MULTIFRACTAL CHARACTERIZATIONS OF INTERMITTENCY IN NONSTATIONARY GEOPHYSICAL SIGNALS AND FIELDS

A MODEL-BASED PERSPECTIVE ON ERGODICITY ISSUES ILLUSTRATED WITH CLOUD DATA

ANTHONY B. DAVIS,† ALEXANDER L. MARSHAK,† WARREN J. WISCOMBE, AND ROBERT F. CAHALAN

NASA – Goddard Space Flight Center, Climate and Radiation Branch (Code 913), Greenbelt, MD 20771, USA.
E-mail: davis@climate.gsfc.nasa.gov

ABSTRACT

There are many reasons for wanting to quantify spatio-temporal correlations in geophysical signals $f(x)$ over a large range of scales r. Standard approaches use either the autocorrelation function $\langle f(x+r)f(x)\rangle$ or the related 2nd-order structure function $\langle[f(x+r)-f(x)]^2\rangle$, equivalently (Wiener-Khinchin theorem), the wavenumber spectrum $E(k)$, with $k \approx 1/r$ as the scale parameter. These are all 2nd-order statistics however, and they do not discriminate well between fields with sometimes radically different spatial properties. For example, in seismic signals the background (possibly instrumental) noise can be modeled as white whereas the interesting events are more like Dirac δ-functions: both components are δ-correlated in the sense of $\langle f(x+r)f(x)\rangle$ and have correspondingly flat wavenumber spectra $E(k) \equiv$ *constant*. In another instance, temporal fluctuations of air temperature are Brownian motion-like, with $\langle[f(x+r)-f(x)]^2\rangle \propto r$ (hence $E(k) \propto k^{-2}$), under quiescent meteorological conditions; unfortunately, the same spectrum is assigned to the occurrence of a quasi-discontinuity marking the passage of a front, as approximated by a Heaviside step function. The issue at hand is resolved by introducing the notion of "intermittency," a concept borrowed from turbulence theory that describes the occurrence of bursts of intense events; statistically speaking, we are faced with the break-down of the prevailing Gaussian paradigm in data analysis. To characterize intermittency, some form of wavelet-type time/frequency (or position/scale) analysis is required. Multifractal approaches to position/scale analysis are particularly easy to exploit: they use higher-order moments as a simple way of sorting the continuum of weak, intermediate, and strong events, and we look for power-law regimes in the resulting scale-dependent statistical quantities at all orders. The two main categories of multifractal analysis, qth-order structure functions and singularity analysis, are surveyed and illustrated with both models and cloud-related data in 1D and 2D. We address in detail the sampling (or "ergodicity") problems that arise as soon as Gaussian assumptions are relaxed and their relation to both stationarity and intermittency is discussed. Finally, we outline how multiscaling has helped to further the theory of cloud-radiation interaction, as applied to the forcing of the climate system and the remote sensing of cloud properties.

---

†Also: Science Systems and Applications, Inc. (SSAI), 5900 Princess Garden Parkway, Lanham, Md.

1. Introduction

1.1 Background

There is an increasing need in the geophysical community for statistical analysis of data. This need is traditionally met with techniques developed within entirely different areas of research. For example, first and second order statistics —means, variances and covariances— found their foremost applications in psychometrics. These well-known quantities are the parameters of the most general multivariate Gaussian distribution. Their counterparts in time-series analysis are the 2-point autocorrelation function and the wavenumber spectrum which proved to be powerful tools for solving engineering problems in communications and signal processing.

Over the decades, it became apparent that Gaussian —and otherwise "thin-tailed"— statistics were ill-suited to describe many random signals that reflect the variability of geophysical fields. Stretched exponentials, log-normal and even power-law distributions were introduced to describe seismic activity, rain-rates, atmospheric turbulence, and numerous other natural phenomena. Opening the Pandora's box of non-Gaussian statistics with "long-" or "fat-tailed" distributions (Waymire and Gupta 1981) raises important questions about sampling. We will refer to these issues generically as "ergodicity" problems, a terminology that better reflects our model-based investigation. The basic question is: *Do space/time averages converge to well-defined ensemble counterparts with sample size and how fast?* More formally put: *How much of probability space do we need to explore to characterize a distribution?* For exactly Gaussian processes, the answer to these questions is more-or-less contained in the "3σ" rule: events more than three standard deviations away from the mean are improbable at the level 0.001.[1] We will show further on examples of natural variability where the standard deviation itself is not even pinned down after many thousands of observations.

In our experience with geophysical data analysis, we have encountered at least two sources of ergodicity problems that generally appear compounded: "intermittency," and "nonstationarity." We are adopting vocabulary from time-series analysis here for simplicity but have either space or time in mind. [Rather than "(non)stationarity," the technically correct usage in random field theory is statistical "(in)homogeneity." However, in keep with cloud-modeling usage, we will reserve this last expression to designate (non-)constant fields which, in turn, are "(non-)trivial" from the times-series perspective.]

- It is natural to think that estimates of statistical properties of geophysical signals are independent of the instant when observations start and stop.[2] This is effectively a "**stationarity**" assumption, i.e., averages are statistically invariant under translations in space or time. Observation time is irrelevant to systems in some kind of dynamical equilibrium. However, very long-range correlations can and do occur in geophysical signals because of the sheer size of the system and the coherent, long-lived structures generated by the large-scale forcing and the highly nonlinear character of the dynamics. In other words, stationarity is generally *not* a good assumption, at least at close range (small scales). The weaker assumption of nonstationarity with stationary increments is generally good enough at all scales (examples to follow). Only when the record is exceptionally long, do we observe a transition to stationarity *per se*. In summary, the questions to ask is *'for how long are the data correlated?'* or *'how long must we wait to isolate independent samples in a given datastream?'*
- Assuming that the nonstationarity has been tamed by focusing on the appropriate quantities to be averaged (e.g., increments), we can still be faced with "**intermittency**" problems. We borrow this concept from turbulence theory to describe the occurrence of sudden bursts of intense variability, very uncharacteristic of Gaussian processes —stationary or not. Following a now well-established tradition in turbulence studies (Parisi and Frisch 1985, Meneveau and Sreenivasan 1987*a*, etc.), we can use "multifractal" statistics to describe intermittency in natural signals. These are straightforward generalizations of the 2nd order statistics mentioned above where moments of all orders —within limits set by sampling considerations— are computed on a scale-by-scale basis and where the dependence on scale is parameterized by power laws. We thus define "scaling" regimes and associated families of exponents.

A priori, nonstationarity and intermittency are purely qualitative attributes when it comes to data analysis. Furthermore, some if not all of the positional information needed to make a statement about stationarity is lost due to the spatial averaging that produces the statistics in the first place. We will show that, in the framework of scale-invariant processes, nonstationarity and intermittency can both be not only detected (cf. §5.1 and §6.3) but precisely quantified as well (cf. §4.5).

Fractal (single-moment) and multifractal statistics were originally perceived as abstract and were criticized for having little bearing on the underlying physics. Serious efforts have been put forth to make multifractal concepts attractive to a broad range of geophysicists (e.g., Davis *et al.* 1994*a*), and their connections with wavelet analysis are now well-understood (Muzy *et al.* 1994). It is true that fractal concepts become mathematically precise in the small-scale limit ... which is generally unjustified on physical grounds. Nevertheless, "physical" fractals (with well-defined inner- and outer-scales) have proven to be very helpful models of reality in a broad range of applications. In fact, the limits of the scaling regimes themselves convey as interesting information on the system as do the exponents, if not more. We will survey our findings in this area with respect to cloud structure and cloud-radiation interaction.

1.2 Overview: The "Laboratory" Model for Geophysical Data Analysis

The paper is organized as follows —in the spirit of a report on a laboratory experiment (cf. Fig. 1). The goal of the "experiment" is to characterize the structure of

[1] Every "Gaussian-type" or "thin-tailed" distribution has a similar rule; e.g., deviations in excess of ≈4.9 σ's from the mean μ are unlikely at the level 10^{-3} for Laplace's probability law:
$$\text{Prob}\{X \le \xi < X+dX\} = \exp[-\sqrt{2}|X-\mu|/\sigma]dX/(\sqrt{2}\sigma).$$

[2] We assume here that all the observations belong to a well-defined class within which we can perform meaningful statistics. At least in atmospheric applications, this may impose external limits on the time of day, the season and the position on the globe.

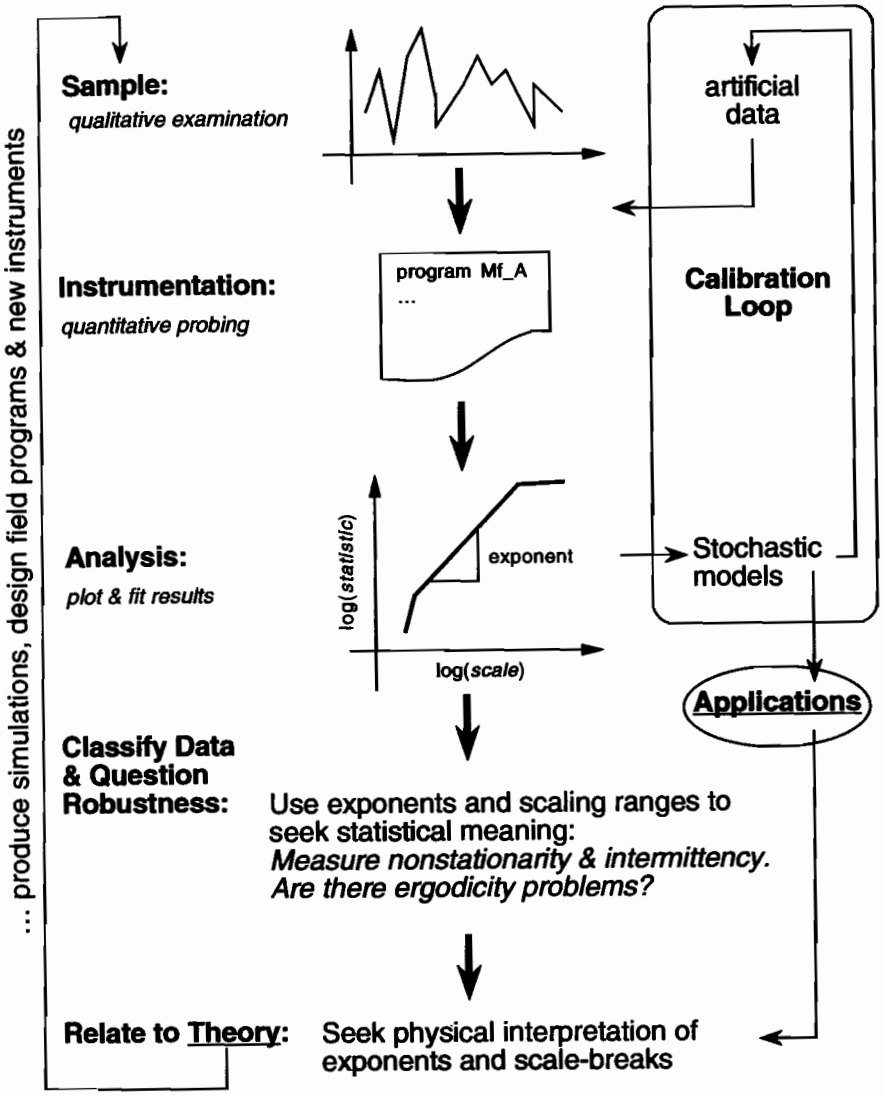


Figure 1: *Flow-Chart for the "Laboratory" Model for Scale-by-Scale Statistical Analysis of Geophysical Data.* There are two feed-back loops. One signifies that new theory makes new predictions to be verified with new data (from new instruments if necessary). The other represents the production of synthetic data with stochastic models to calibrate the statistical "instrumentation" with standard input. Stocahstic modeling also feeds into the body of theory that explains the data. In the case of cloud-radiation theory, the numerical modeler can control the weather in his digital cloud system, the simplest models may be amenable to analytic methods.

marine stratocumulus for the purposes of radiative transfer computations. These are climatologically important cloud systems, and our main motivations to study them are (*i*) improved models for understanding the atmospheric radiation budget, hence present and future climate, and (*ii*) improved methods for retrieving cloud properties from remotely-sensed signals.

- Section 2 ("Sample Collection"): we present 1D and 2D data pertaining to marine stratocumulus and further motivate their study.
- Section 3 ("Instrumentation") describes the basic tools of scale-by-scale statistical analysis: coarse-graining, autocorrelation, structure functions and spectral analysis.
- Section 4 ("Results") establishes the relevance of power-law parameterizations for scale-dependent cloud statistics and defines notations for the associated exponents.
- Section 5 ("Semi-Empirical Criteria") shows how we interpret certain scaling properties, with an emphasis on ergodicity issues: stationarity, intermittency and the onset of sampling problems are discussed.
- Section 6 ("Theoretical Considerations"): we rephrase our outlook on data analysis in terms of symmetry and broken symmetry; this impacts directly our understanding of cloud structure and how it transpires in satellite images. We also draw parallels between our methods and those of statistical physics.
- Section 7 is a summary.
- Appendix ("Calibration and Simulation"): a number of scale-invariant models are introduced and classified according to the criteria in Section 5. These algorithms for generating synthetic data have many applications, our present concerns being (1) assessment of the reliability of analysis procedures and (2) simulation of realistic clouds-in-a-computer that enable numerical radiation transport studies. In all cases, the statistical properties —namely, the exponents— are known *a priori*.

2. SAMPLE COLLECTION (Cloud Data as an Object of Study)

When designing an experiment, generally to test some hypothesis, the first questions to ask are *'What are we going to study?'* and *'Why?'*. We can view data analysis as a straightforward experimental procedure where the object of study is the data itself (Fig. 1). We will assume it to be stored in a 1D or 2D array of real numbers residing in computer memory. In this section, we present geophysical data and present a rationale for an in-depth study of its statistical properties. The next questions are closely related: to each other: *'What properties are we interested in?'* and *'What instruments will we use to probe our sample?'*. They are addressed in sections 3–4.

We present here data used in our specific research area: internal structure of marine stratocumulus (Sc) and its impact on radiation transport. Generally speaking, cloud-radiation interaction is a source of considerable uncertainty in the prediction of climate and climate change. Being both persistent and extended, marine Sc layers are responsible for a large portion of the Earth's global albedo, hence the planet's overall energetic equilibrium. A robust statistical characterization of marine Sc structure is therefore in order. In particular, this will allow us to develop statistically realistic cloud models

which can in turn be used to investigate radiation issues. An improved understanding of how physical cloud properties relate to their radiation fields has important spin-offs in the area of remote sensing. This is the only cost-effective way of monitoring cloud cover from synoptic to pixel scales (several kilometers or meters, depending on the device).

2.1 One-Dimensional in Situ Transects of Liquid Water Content from FIRE

2.1.1 Internal Cloud Structure Using Taylor's Frozen Turbulence Hypothesis

Clearly, there is no better way to study cloud structure than by direct probing. This calls for a fully instrumented aircraft and, because of the costs involved, few datasets of this type are available. Furthermore, cloud liquid water content (LWC) measurement is still an area of active research (Gerber *et al.* 1994). We present here transects of LWC in marine Sc that we will use to illustrate 1D data analysis in the remainder of the paper.

In Figs. 2*a-e*, we show representative samples of LWC vs. time from five flights (or flight legs) during FIRE[1] in June–July 1987 off the coast of southern California. Following a well-established practice in the turbulence literature (Taylor's frozen turbulence hypothesis), we perceive these time-series as 1D cuts through the spatially variable LWC field:

$$f_i(x_m), x_m = m\ell, m = 1,\ldots,M_i\ (i = 1,\ldots,5). \qquad (1)$$

Table 1 shows the important parameters of the datasets, described in more detail by Davis *et al.* (1996*a*). In particular, they tentatively relate the down-spikes that characterize Figs. 2*a,b* to dynamical instabilities, and they question the reality of the strong "dip" in Fig. 2*c*. In the following analyses, events affected by this feature are ignored but Marshak *et al.* (1996) examine the consequences of not eliminating the spurious dip.

Table 1: *FIRE Liquid Water Content (LWC) Database.* The statistically relevant parameters of the various datasets are collated. They were obtained from an airborne platform during the FIRE 1987 stratocumulus experiment, off San Diego (Ca.). A nominal aircraft speed of 100 m/s was used to convert time to space, the sampling rate being 20 Hz.

Date (1987)	Time (GMT)	M (points)	$\log_2 M$	Length L (km)	Character	Comment
6/30	22:41	28672	14.81	143	spiky	downward
7/02	02:23	16384	14	82	spiky	downward
7/14	23:09	65536	16	328	smooth	suspicious dip
7/16	17:17	8192	13	41	mixed	spikes down
7/16	18:19	12020	13.55	60	mixed	spikes up

2.1.2 Visualizing Intermittency with Small-Scale Absolute Gradients

The most interesting (i.e., strong and somehow "organized") features in Figs. 2*a–e* are the large and well-localized downwards deviations that occur intermittently but

[1] First ISCCP Regional Experiment (ISCCP = International Satellite Cloud Climatology Project).

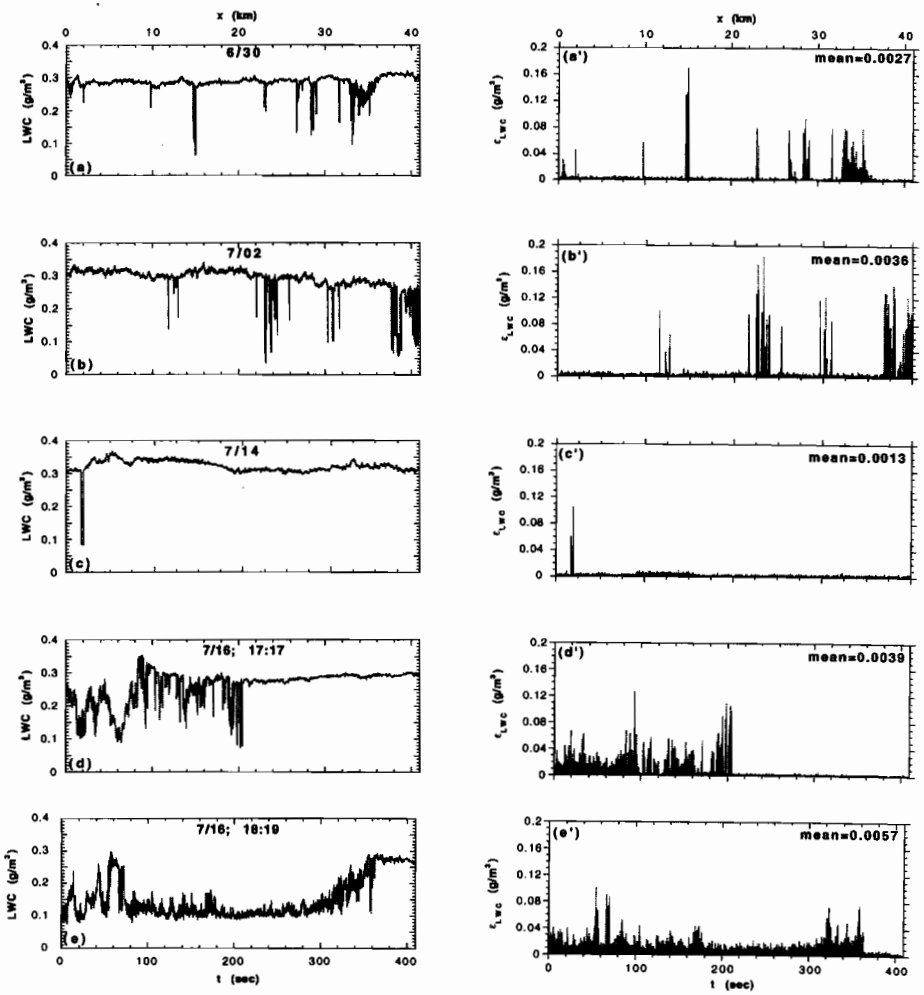


Figure 2: *Liquid Water Content (LWC) Transects from In Situ Airborne Probings of Marine Stratocumulus.* (*a–e*) Representative portions of the five datasets described in Table 1; these are examples of nonstationary processes with stationary increments (Fig. 4 and §5.1) and multiscaling structure functions (Fig. 6 and §4.2). These data were collected in marine stratocumulus during FIRE in 1987 off the coast of southern California. (*a'–e'*) Absolute next-neighbor differences for the data in panels (*a–e*); these are examples of intermittent stationary processes (Fig. 4 and §5.1) with *bone fide* multifractality (Table 2 and §5.3), as revealed by singularity analysis (Fig. 8 and §4.4).

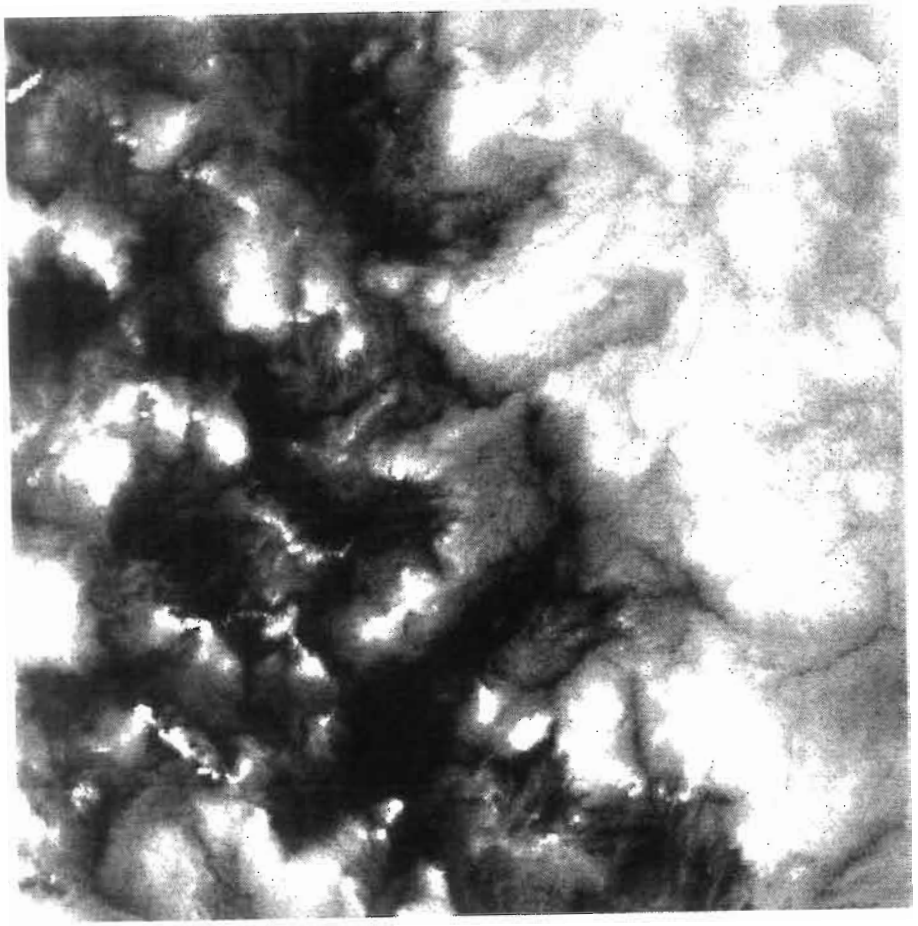


Figure 3: *Radiance Field of Marine Stratocumulus.* Gray-scale rendering of a 4096×4096 portion of a Landsat image of a typical cloud deck off the coast of southern California. This scene was captured at visible wavelengths (channel 2 of the Thematic Mapper) on June 30 1987; so the climatological conditions are similar to those prevailing when the *in situ* LWC data in Fig. 2 were obtained.

nevertheless seem to cluster. The large jumps that characterize these events are highlighted by taking the absolute gradients of the data at some small (but presently unspecified) scale η. Adopting units where the sampling scale (grid constant) $\ell = 1$ for simplicity, we have:

$$\varepsilon_{ij}(\eta;x_m) = |f_i(x_{m+\eta})-f_i(x_m)|,\ m = j,\ldots,M_i-\eta+j\ (j = 1,\ldots,\eta;\ i = 1,\ldots,5). \qquad (2)$$

The new data obtained from that in Figs. 2*a-e* with $\eta = \ell = 1$ are presented in Figs. 2*a'-e'*.

In turbulence studies where $f(x_m)$ is most often velocity, η is taken to be the "Kolmogorov" scale where dissipation forces start to dominate inertia. Furthermore, it is traditional in this context to take squares rather than absolute values in Eq. (2) since this yields a 1D cut through the physically important field that describes the local rate of kinetic energy dissipation (Meneveau and Sreenivasan 1987*a*).

2.2 A Two-Dimensional Radiance Field from LANDSAT Captured During FIRE

An attractive alternative to *in situ* probing of cloud structure is to use high-resolution satellite imagery. It is relatively inexpensive compared to outfitting and flying research aircraft, more comprehensive than aircraft probing or ground-based radiometry (being 2D), and more frequent in time.

Figure 3 is a gray-scale rendering of a large (4096×4096) section of a cloudy LANDSAT scene:

$$f(x_{m_1},y_{m_2}),\ (x_{m_1},y_{m_2}) = (m_1,m_2)\ell,\ m_1 = 1,\ldots,2^n,\ m_2 = 1,\ldots,2^n\ (n = 12). \qquad (3)$$

The signal f is digitized over 256 levels and is almost proportional to nadir-viewing radiance at satellite level (≈800 km). LANDSAT's radiometer was not originally designed for such bright targets as clouds, so saturation (at $f = 255$) occurs frequently, 17% in our case. In order to avoid spurious saturation effects, the up-coming statistical analyses use only the 2048×4096 leftmost pixels which are only 7% saturated.

3. INSTRUMENTATION (Scale-by-Scale Analysis Tools for 1D or 2D Datasets)

Pursuing our analogy between data analysis and laboratory work, we describe the first part of the experimental procedure. *What "measurements" are we going to do? What "instruments" are we going to use?* The object under scrutiny is cloud data stored in a large portion of computer (and possibly peripheral) memory. The instruments are programs that process this data; their output constitute statistical measurements. This new "data," residing in far less memory, describes partially the dataset. In essence, we are observing the statistical "behavior" of the data/subject with different devices and, in a sense we will define in §3.3, under different "experimental conditions."

All our instruments have two computational stages, performed in sequence (§3.1) or in parallel (§§3.2–4). First comes an analysis procedure that yields, in general, a quite large number of random variables by resampling and operating on the data. This is followed by a spatial/ensemble averaging. Consider, as an example, the computation of

1-point variance $\langle f^2\rangle-\langle f\rangle^2$: it calls for (1) forming the 1st and 2nd powers of $f(x)$ then (2) obtaining their averages over all the available data. In the following, we will consider exclusively 2- and more-point statistics that contain information about correlations (or "structure") in the datastream. Since we always compute statistical properties at a specific scale r, we refer to these techniques collectively as "scale-by-scale" analyses.

3.1 Spectral Analysis

Correlations in random data can be studied via Fourier analysis, leading to energy spectrum estimation. This is a well-traveled approach to scale-by-scale analysis where the scale parameter is wavenumber k, related to the length scale $r = 1/k$ in physical space.

Let $\tilde{f}(\boldsymbol{k})$ be the d-dimensional Fourier transform of a field $f(\boldsymbol{x})$ defined on $[0,L)^d$:

$$\tilde{f}(\boldsymbol{k}) = \int_{[0,L)^d} f(\boldsymbol{x}) \exp(2\pi i \boldsymbol{x}\boldsymbol{k})\, \mathrm{d}^d\boldsymbol{x}\ , \tag{4a}$$

where the normalized wavevector $\boldsymbol{k}L$ scans $\mathbb{Z}^2$. In the discrete case, the above integrals become sums (with $\mathrm{d}^d\boldsymbol{x} = \ell^d$):

$$\tilde{f}(\boldsymbol{k}) = \ell^d \sum_{\boldsymbol{x}} f(\boldsymbol{x}) \times \exp(2\pi i \boldsymbol{x}\boldsymbol{k}). \tag{4b}$$

For $d = 1$, k now goes from $-k_\mathrm{N}$, excluded, to $+k_\mathrm{N}$, included, where

$$k_\mathrm{N} = \frac{1}{2\ell} \tag{5}$$

is the maximal (or "Nyquist") wavenumber, by steps of $\Delta k = 1/L_2$ with[1]

$$L_2 = M_2\ell = 2^{[\log_2 M]}\ell, \tag{6}$$

$[\cdot]$ designating integer part. For $d = 2$, the subset $(-k_\mathrm{N}L_2,+k_\mathrm{N}L_2]\otimes(-k_\mathrm{N}L_2,+k_\mathrm{N}L_2]$ of $\mathbb{Z}^2$ is scanned by $\boldsymbol{k}L_2 = (k_xL_2,k_yL_2)$. We note that, for $f(\boldsymbol{x}) \in \Re$, $\tilde{f}(-\boldsymbol{k}) = \tilde{f}(+\boldsymbol{k})^*$.

In Fourier space, the spatial averaging step is replaced by a summation over phases, equivalently, over wavenumber sign. We compute the energy[2] spectrum in $d = 1$ from

$$E(k) = \langle \|\tilde{f}(+k)\|^2 + \|\tilde{f}(-k)\|^2\rangle \Delta k = \frac{2-\delta_{0k}}{L} \langle \|\tilde{f}(k)\|^2\rangle \tag{7a}$$

for $kL \in \mathbb{N}$, and where $\delta_{0k} = 1$ for $k = 0$, 0 otherwise. In the discrete ($\ell > 0$) case, the Fourier series is truncated, hence

[1] The most popular Fast Fourier Transform packages (e.g., Press *et al.* 1993) require M to be a power of 2. If this is not the case, the first and last M_2 data points can be treated as two realizations in an ensemble average. Generally speaking, ensemble averaging over a number of datasets poses no special problem as long as they have the same ℓ and L_2; otherwise, common units for k and $E(k)$ must be defined if different L_2's occur and k-bins must be used if different ℓ's occur.

[2] $E(k)$ is called interchangeably "power" or "energy" or "wavenumber" or "frequency spectrum," and sometimes "periodogram."

$$E(k) = \frac{2-\delta_{0k}}{L_2} \langle \|\tilde{f}(k)\|^2\rangle \tag{7b}$$

for $kL_2 = 0,1,\ldots,L_2/(2\ell)$. In $d = 2$ cases, we assume statistical isotropy and sum $\langle \|\tilde{f}(k)\|^2\rangle$ over circles of radius $k = |\boldsymbol{k}| = \sqrt{k_x^2+k_y^2}$ in Fourier space. For continuous spectra ($\ell = 0$, $L \to \infty$, hence $\Delta k \to 0$), we have

$$E(k) = \int_{|\boldsymbol{k}|=k} \langle \|\tilde{f}(k)\|^2\rangle\, \mathrm{d}^2\boldsymbol{k} = k\int_0^{2\pi} \langle \|\tilde{f}(k\cos\theta,k\sin\theta)\|^2\rangle\, \mathrm{d}\theta \tag{8a}$$

for $k \geq 0$. In the discrete ($L < \infty$, $\ell > 0$) case, kL_2 goes from 0 to $k_{\max}L_2 = [\sqrt{2}k_\mathrm{N}L_2]$ and

$$E(k) = \frac{1}{L_2} \sum_{kL_2 \leq |\boldsymbol{k}|L_2 < kL_2+1} \langle \|\tilde{f}(k)\|^2\rangle \tag{8b}$$

for $kL_2 = 0,1,\ldots,k_{\max}L_2$.

It is sometimes advantageous to cumulate and average $E(k)$ and k in octave-wide bins (i.e., by factors of 2 in k) for $k > 0$ (Davis *et al.* 1996*a*). There are precisely $[\log_2 M]-1$ bins when $d = 1$: $\{kL_2 \in \mathbb{N};\ 2^i \leq kL_2 < 2^{i+1}-1\}$, for $i = 0, \ldots, [\log_2 M]-2$; this excludes the Nyquist frequency, the most aliased anyway (Press *et al.* 1993). In $d = 2$ cases, one more bin can be populated by wavevectors with their modulus between k_N and $k_{\max}$: in all, $\{\boldsymbol{k}L_2 \in (-k_\mathrm{N}L_2,k_\mathrm{N}L_2]^2;\ 2^i \leq |\boldsymbol{k}|L_2 < 2^{i+1}-1\}$ for $i = 0, \ldots, [\log_2 M]-1$.

3.2 Two-Point Correlation Analyses at Order 2

There are two physical space counterparts of spectral analysis for correlation studies. First, one can form the products $\delta f(x+r)\delta f(x)$ of the fluctuating part $\delta f(x) = f(x)-\langle f\rangle$ of a 1D signal at two points and spatially (then, if necessary, ensemble-) average them to obtain the autocorrelation function:

$$\langle \delta f(x+r)\delta f(x)\rangle = \langle f(x+r)f(x)\rangle - \langle f\rangle^2. \tag{9a}$$

This well-known 2nd order 2-point statistic does not give us any new information since it is related to the energy spectrum in Eqs. (7–8) by the Wiener-Khinchin (W–K) relation:

$$\langle f(x+r)f(x)\rangle - \langle f\rangle^2 = \int_0^\infty \cos(2\pi rk)E(k)\mathrm{d}k. \tag{9b}$$

At $r = 0$, the l.h.s. reduces to the 1-point variance, and the r.h.s. is the integral of $E(k)$; so $E(k)\Delta k$ is simply the part of the variance that comes from scales $\approx 1/k$.

One can also form "increments,"

$$\Delta f(r;x) = f(x+r)-f(x), \tag{10}$$

in 1D and compute the 2nd-order "structure function,"

$$\langle \Delta f(r;x)^2\rangle = \langle [f(x+r)-f(x)]^2\rangle, \tag{11a}$$

also known as a "semi-variogram" (Christakos 1992). Here again there is a W–K relation with the energy spectrum:

$$\langle \Delta f(r;x)^2 \rangle = 2 \int_0^\infty [1-\cos(2\pi r k)] E(k) \mathrm{d}k, \tag{11b}$$

as results from identities $\langle \Delta f(r;x)^2 \rangle \equiv 2[\langle f^2 \rangle - \langle f(x+r) f(x) \rangle] \equiv 2[(\langle f^2 \rangle - \langle f \rangle^2) - \langle \delta f(x+r) \delta f(x) \rangle]$. Theoretically (i.e., when doing ensemble rather than spatial averages), these operations are only meaningful in "broad-sense" stationary situations where $\langle \delta f(x+r) \delta f(x) \rangle$ depends only on r. However, Eq. (11*b*) generalizes to nonstationary signals with (broad-sense) stationary increments, namely, where $\langle \Delta f(r;x)^2 \rangle$ is function of r alone.

In isotropic 2D situations, $\delta f(x+ru) \delta f(x)$ or $[f(x+ru)-f(x)]^2$ can be obtained by averaging over the allowable domain of x and the orientation of the unit vector u. This straightforward approach quickly becomes computationally intractable since it requires $\sim N^2$ operations (where $N = M^2$ is the total number of points). In contrast, FFT implementations of Eqs. (8*b*) or (11*b*) require only $\sim N \ln N$ operations; in this case however, we interpret $E(k)$ as the so-called "1D" spectrum obtained by dividing the r.h.s. of Eq. (8*b*) by $\sum_{k L_2 \le k L_2 < k L_2 + 1} 1$, i.e., averaging rather than just summing $|\hat{f}(k)|^2$. Another approach (with only $\sim N$ operations and generalizable to higher orders) is adopted in the remainder of this study: to treat rows and columns as an ensemble of 1D datasets.[1]

3.3 *q*th-Order Structure Functions

How can one gain *new* information in the framework of 2-point statistics? Simply by looking at moments of order $q \neq 2$. The random variables of interest are then $|\Delta f(r;x)|^q$ and averaging yields the qth-order structure function:[2]

$$\langle |\Delta f(r;x)|^q \rangle = \langle |f(x+r)-f(x)|^q \rangle. \tag{12}$$

Unfortunately, we lose the W–K connection and the computational efficiency of FFTs in 2 or more dimensions. However, the focus on increments is akin to a high-pass filtering. Therefore, at the cost of using the $\sim N$ coefficients of a discrete wavelet decomposition of $f(x)$ as surrogates for $\Delta f(r;x)$, the utilization of multiresolution analysis (Mallat 1989) will lead to efficient computational algorithms.

What insight do we gain by varying the parameter q? Of all possible values that the increment $\delta = |\Delta f(r;x)|$ can take, we can identify

- "typical" values that occur most frequently, being near the mode of the pdf $p_r(\delta)$,
- "mean" values that dominate the average for $q = 1$ (i.e., maximize $\delta p_r(\delta)$),
- "r.m.s." values that dominate the average for $q = 2$ (i.e., maximize $\delta^2 p_r(\delta)$).

There are also ever larger and rarer values that dominate higher order statistical moments: $q = 3, 4$, etc. So, increasing q amounts to looking at the more extreme values of $|\Delta f(r;x)|$.

[1] In theory, sampling is poor ($\sim N$ out of $\sim N^2$ possible events); but $N = M_x \times M_y$ is already large in general.

[2] Validation of structure function computation and sensitivity studies with respect to amount of data can be performed with the help of the models presented in the Appendix, more specifically in §A.2 and §A.4.

This is akin to changing the experimental conditions (e.g., temperature T) and observing changes in the state of a macroscopic physical system made of many macroscopic elements interacting with each other; for instance, a real gas or a ferromagnet. By lowering T the system can be forced into otherwise very unlikely configurations. The observable (macroscopic) changes can be very subtle or extremely dramatic, like when a phase transition occurs: e.g., a mole of H_2O molecules goes from vapor to liquid, to solid. Whatever the outcome, we learn more about the specificity of the system (data) by exploring as large a range of T's (q's) as possible.

3.4 *Running Averages and the qth-Order Moments of the Coarse-Grained Field*

The high-pass filtering implicit when taking increments over various scales eliminates the large-scale mean value $\langle f \rangle$ from the picture. What about the converse operation, computing local means (low-pass filter output) at various scales?

Consider for instance the running mean[1] of $f(x)$ over $[x,x+r)$:

$$\mu_f(r;x) = \frac{1}{r} \int_x^{x+r} f(x') \mathrm{d}x', \tag{13a}$$

for $x = [0,L-r)$ and $r > 0$. In the 2D case, spatial averaging is over the square domain $[x,x+r) \otimes [y,y+r)$ of area r^2. Consider also running variance of $f(x)$ over $[x,x+r)$:

$$\sigma_f^2(r;x) = \frac{1}{r} \int_x^{x+r} f(x')^2 \mathrm{d}x' - \mu_f(r;x)^2. \tag{13b}$$

In one application of the above running means and variances, we hold x constant and vary the scale r continuously until some kind of "convergence" is obtained (cf. discussion of Fig. 8*a*). Instead, we can think of r as a fixed parameter; then $\mu_f(r;x)$ and $\sigma_f(r;x)$, obtained in Eqs. (13*a*,*b*), are random numbers. From this perspective, we can study their statistical properties by averaging over x, the position of the segment or square (cf. discussion of Fig. 8*b*), as well as seek the correlations[2] between $\mu_f(r;x)$ and $\sigma_f(r;x)$.

By resampling x at intervals of length r in Eq. (13*a*), we obtain a "coarse-grained" version of the original field $f(x)$, with L/r pixels rather than L/ℓ. This is of particular interest for the absolute gradients defined in Eq. (2). In the 1D continuum limit used in Eqs. (13*a*), we define

$$\varepsilon(r;x) = \mu_\varepsilon(r;x) = \frac{1}{r} \int_x^{x+r} \varepsilon(x') \mathrm{d}x', \; r \ge 0, \tag{14a}$$

[1] In the language of continuous wavelet transforms, Eq. (13*a*) is the projection of the signal $f(x)$ onto a functional space of "scaling functions." In this case, the scaling function equals $1/r^d$ on a compact support of measure r^d and 0 elsewhere; Gaussian and otherwise variable "windows" have also been considered.

[2] This is the basis of the "spatial coherence" method developed by Coakley and Bretherton (1982) for recovering fractional cloudiness from satellite radiances.

and similarly in 2D. This translates to

$$\varepsilon(r;x) = \frac{1}{r} \sum_{x'=x}^{x+r-1} \varepsilon(1;x'),\ r \geq 1, \qquad (14b)$$

in the discrete case where the small-scale $\varepsilon(1;x)$ field is defined on a grid with constant $\ell = 1$; in other words, units of length where $\eta = 1$ are employed in Eq. (2) when used to obtain $\varepsilon(1;x)$ from $f(x)$. Generalization from 1D to 2D is again straightforward and, as for the wavelet-based surrogates for the $\Delta f(r;x)$'s mentioned in §3.2, efficient computation of $\varepsilon(r;x)$ at appropriately selected x's can be implemented in the framework of multiresolution analysis (cf. graphics by Davis *et al.* (1994*a*)).

Having obtained the non-negative random numbers in Eqs. (14*a*,*b*), we can take their qth powers and average the results over x, the positions of segments or squares, to obtain[1] $\langle \varepsilon(r;x)^q \rangle$ with the same advantages in terms of experimental "temperature" control as discussed earlier for $\langle |\Delta f(r;x)|^q \rangle$.

4. RESULTS (Power-Law Regimes and Scale-Breaks)

In this section, we examine the output of our laboratory "instruments." These "measurements" are a new —and highly compressed— form of data: statistical quantities, always parameterized by scale. Typically, the scale parameter r or k spans a large range of values and, in general, the statistics (e.g., $E(k)$, $\langle |f(x+r)-f(x)|^q \rangle$, or $\langle \varepsilon(r;x)^q \rangle$) do too. The usual way of visualizing quantities with large ranges is to use log-log plots and we naturally ask: *Are there significant ranges of scale where* log(*statistic*) *is linear in* log(r)*?* This is often the case and they are called "scaling" regimes. In such regimes the important parameter is the slope on the log-log plot, equivalently, the exponent in the associated power law. In this section, we introduce notations[2] for a number of exponents, present results for our cloud-related test data and discuss their most striking features.

4.1 Scaling in Spectral Analysis

Figure 4*a* shows octave-binned energy spectra for the five 1D datasets in Table 1, partially illustrated in Figs. 2*a–e*. Log-log axes are used and the k-ranges are different due to the different lengths of the datasets (L_2's). We see good scaling in all cases, at least for $k_N/10^2 \lesssim k \lesssim k_N$ (last 7 octaves), with good agreement in the prefactors for 4 of the datasets. The dispersion at small k (large scales) reflects the visual diversity of Figs. 2*a-e*. The odd dataset is also the longest ($L \approx 330$ km, about a half of all the FIRE LWC data), only an eighth of which is illustrated in Fig. 2*c*. Apart from the suspicious dip visible in that figure (but not incorporated in our analyses), this data looks very smooth. Spectrally, this translates in too ways: (1) the prefactor in the scaling regime is

[1] Validation of these computations and sensitivity studies with respect to amount of data can be performed with the help of models presented in the Appendix, specifically in §A.3.

[2] Unfortunately, there are several co-existing standards in the literature.

significantly smaller than for the four other datasets; (2) there is a specific scale $1/k$ at which the Fourier modes stop increasing with scale and become constant. This is known as an "integral" (correlation) scale, denoted R; we estimate $R \approx$ 20–40 km.

Figure 4*b* shows our results for octave-binned $E(k)$ for the five absolute next-neighbor gradient fields associated with the LWC data analyzed in Fig. 4*a*. We notice that the behavior is somewhat more erratic than in Fig. 4*a* but still shows a reasonable correlation between log$E(k)$ and log$k$ for at least two decades in scale (≈7 octaves).

Figure 4*c* shows the octave-binned estimates of the ensemble-average spectra for the LWC fields ($f_i(x)$, $i = 1,\ldots,5$) and for their absolute small-scale gradients $\varepsilon_i(\ell;x)$, from Eq. (2) with $\eta = \ell$ (in this special case, the second subscript is unnecessary). For both spectra, we see power-law behavior over no longer two but three decades (10 octaves) in scale. In other words, the realization-to-realization variability in $E(k)$ at large scales has been damped by the ensemble-averaging. Unfortunately,[1] the break at $R \approx$ 20–40 km is not apparent in the average data because it starts at $k = 1/\min\{L_2\}$ where $\min\{L_2\} \approx 40$ km according to Table 1. However, we are confident that the transition from increasing to constant variance as scale increases (k decreases) will be consistently observed, whether at 20–40 km or more, as longer LWC datasets become available. This is because LWC is bounded on physical grounds: it is non-negative by definition with an upper bound which is a some small portion of the total amount of water vapor that the laws of thermodynamics allow a column of atmosphere to contain. This means that Fourier modes cannot be arbitrarily strong; a cut-off must occur when they reach the climatological mean LWC, otherwise fluctuations overwhelm the dc ($k = 0$) component and the bounds will be exceeded.

The spectral scaling exponent β in

$$E(k) \sim k^{-\beta} \qquad (15)$$

has quite different values for the two types of data: ≈1.4 for the LWC, and ≈0.7 for the associated ε-fields; we discuss the statistical significance of this difference in §5.1.

Figure 5 shows $E(k)$ vs. k in log-log axes for the l.h. half of the Landsat image in Fig. 3. In this case, there is so much averaging involved that octave-binning is not required to reduce the statistical noise. We see a clear-cut break at the integral scale: roughly constant Fourier amplitudes ($\beta \approx 0$) for scales larger than $R \approx 20$ km. There is also a break at ≈0.2 km, a scale we denote η_{rs}. Between $1/R$ and $1/\eta_{rs}$, $E(k)$ follows a power law with $\beta \approx 2.0$ and beyond $1/\eta_{rs}$ a steeper law, with an exponent in excess of 3, the differentiability limit. In other words, small-scale variability is smoother than expected.[2]

Davis *et al.* (1996*b*) investigate the scale break at η_{rs} theoretically, relating it to a radiative smoothing phenomenon mediated by horizontal photon transport by multiple scattering. One outcome of this study is a simple expression for the "radiative smoothing" scale η_{rs} in terms of mean optical and geometrical cloud properties.

[1] Because very long records are needed, it is not always easy to estimate R in large geophysical systems.

[2] At the very smallest scales (≈60 m, twice the pixel scale), another flattening occurs due to digitization.

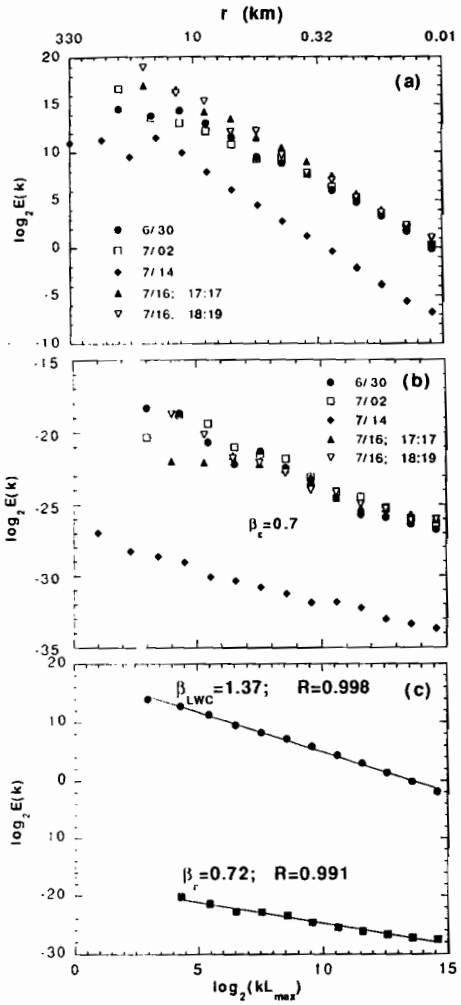


Figure 4: *Energy Spectra of 1D LWC data.* (*a*) Five individual wavenumber spectra for the LWC datasets listed in Table 1, using octave-bins and log-log axes. (*b*) Same as (*a*) but for the absolute small-scale gradient fields. (*c*) Ensemble-average spectra from the data in panels (*a*) and (*b*), weighted by the overall lengths in Table 1.

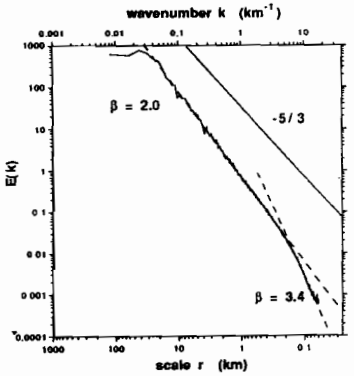


Figure 5: *Energy Spectrum of 2D Radiance data.* Log-log plot of the wavenumber spectrum of the Landsat data in Fig. 3.

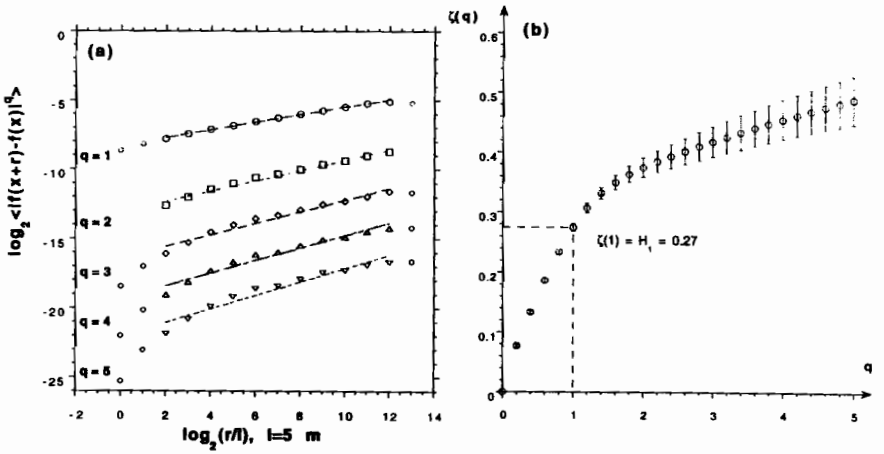


Figure 6: *Structure Functions of 1D LWC data.* (*a*) Scaling for $q = 1(1)5$. (*b*) $\zeta(q)$ exponents.

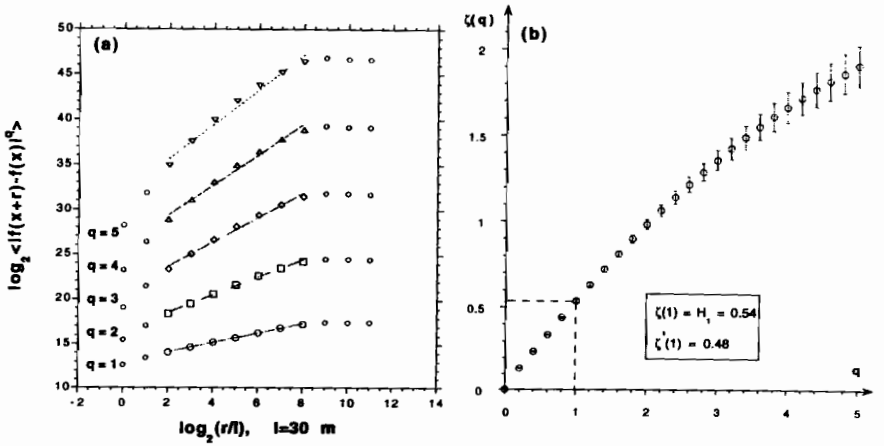


Figure 7: *Structure Functions of 2D Radiance data.* (*a*) Scaling for $q = 1(1)5$. (*b*) $\zeta(q)$ exponents.

4.2 *Multifractal Analysis 1, Higher-Order Structure Functions*

Figure 6*a* shows the ensemble-average $\log\langle|f(x+r)-f(x)|^q\rangle$ vs. $\log r$ for integer-valued $q = 1,\ldots,5$ using the 1D LWC data as input. We note that by taking $r = \ell$, $r = 2\ell$ ($= 1/k_N$), $r = 4\ell$, etc., up to $r = \min_{i=1,\ldots,5}\{L_{2i}\}$, we obtain one more data point than with $E(k=1/r)$, namely, at the pixel scale ℓ. This extra datum enables us to see some evidence of a break in the scaling at $\approx 2\ell$–4ℓ (10–20 m). At any rate, we see the break at the integral scale (i.e., $R \approx 20$ km) better than in the spectral data. Between these two limits, we have power laws for every q. Adopting the notation used in the turbulence literature, we have

$$\langle|f(x+r)-f(x)|^q\rangle \sim r^{\zeta(q)}. \tag{16}$$

Figure 6*b* shows $\zeta(q)$ vs. q for $0 \le q \le 5$. Muzy *et al.* (1993) have developed an elegant method for estimating $\zeta(q)$ directly from the modulus of continuous wavelet transforms.

Figure 7*a* shows the scaling of the qth-order structure functions for the l.h. half of the 2D data in Fig. 3, showing good scaling between η_{rs} and R and where both the scale-breaks, first characterized spectrally in Fig. 5, are apparent. (Here however, we would significantly underestimate the integral scale $R \approx 20$ km by a factor of ≈ 2 due to the saturation at $f(x) = 255$.) Figure 7*b* shows the corresponding $\zeta(q)$ function.

We note the concavity[1] of $\zeta(q)$ and, since $\zeta(0) = 0$ to ensure proper normalization, we can define a non-increasing hierarchy of exponents:

$$H(q) = \frac{\zeta(q)}{q}. \tag{17}$$

Data like ours yielding non-constant $H(q)$, hence a non-linear $\zeta(q)$, is "multifractal" in the sense of Parisi and Frisch (1985); otherwise (i.e., when $\zeta(q) \propto q$), it is "monofractal."

Two values of q are of particular interest. At $q = 2$, we retrieve the scaling for the 2nd order structure function; for power law statistics, the W–K relation in Eq. (11*b*) yields (e.g. Monin and Yaglom 1975, p. 92)

$$\beta = 2H(2)+1 = \zeta(2)+1. \tag{18}$$

Incidentally, relation (18) is well verified numerically by our data, especially if octave-bins are used in the spectral analysis (Davis *et al.* 1996*a*).[2]

At $q = 1$, we can retrieve the fractal dimension D_g of the graph of the data, viewed as a set $g = \{x, f(x)\}$ embedded in $d+1$ dimensions (e.g. Falconer 1990):

$$D_g = (d+1)-H(1) = (d+1)-\zeta(1). \tag{19}$$

[1]This property follows from characteristic function theory (cf. Feller 1971). The Fourier transform of the pdf of a real random variable ξ is its characteristic function $\phi(t) = \langle\exp(it\xi)\rangle$. Non-negativity of the pdf implies that $\phi(iu)$ is real and non-increasing for real u. By the same token, $\ln\phi(t)$, the cumulant-generating function of ξ, is convex for purely imaginary t. Setting $\xi = \ln|f(x+r)-f(x)|$ for fixed r and $t = iq\ln r$, we see from Eq. (16) that $-\zeta(q)$ is proportional to a cumulant-generating function, hence the concavity of $\zeta(q)$.

[2]This constitutes an internal calibration of the instrumentation in our data analysis lab; see Appendix for external calibration procedures using standard input.

We can therefore equate $H(1)$, the "mean Hölder exponent," to the codimension of the graph of the data. This quantity has natural bounds. If $H(1) = 0$, then $D_g = d+1$: the graph fills a finite portion of $\Re^{d+1}$. If $H(1) = 1$, then $D_g = d$: the graph is as smooth as the support $\Re^d$ of x, namely, a Euclidian line or plane.

4.3 *Running Means and Variances*

Figure 8*a* demonstrates the practical importance of the integral scale R —a 2-point statistical construct— for a 1-point statistic which is *a priori* simpler. We have plotted $\mu_f(r;x)$ vs. r for 14 different non-overlapping 40 km sections from the LWC database. Running means begin to stabilize only at $r \approx 20$ km, some to values quite far from the ensemble-mean of 0.29 g/cm^3 (beyond one ensemble-σ). This tells us two things. First, we need a stretch of at least one or two integral correlation lengths before we can even talk about a mean, even locally. Second, this data is highly non-ergodic: no physically attainable length of LWC data seems to be enough to reach a "climatological" average.[1]

Figure 8*b* shows $\sigma_f(r;x)$ vs. r for the same 14 sections, this time in log-log axes. The striking feature is the jumpiness of running variance due to localized events, clearly the intermittently distributed clusters of down-spikes in Figs. 2*a,b,d,e*; in one case, a factor of ≈ 2 is gained even after accumulating for ≈ 30 km. Such jumps are very unlikely in a process obeying Gaussian statistics. The r.m.s. ensemble-average, $\langle\sigma_f^2(r;x)\rangle^{1/2}$, is also plotted (bold dots): for r between 0.15 and 20 km, it follows a power law in $r^{0.2}$. The exponent is numerically identical to $\zeta(2)/2 = (\beta-1)/2$m, not a coincidence: $\langle\sigma_f^2(r;x)\rangle$ should be of the same order of magnitude as $\langle[f(x+r)-f(x)]^2\rangle$. Here again, no convergence to the estimated ensemble-mean σ (0.05 g/cm^3) is in sight, even at $r = 40$ km $\gtrsim R$.

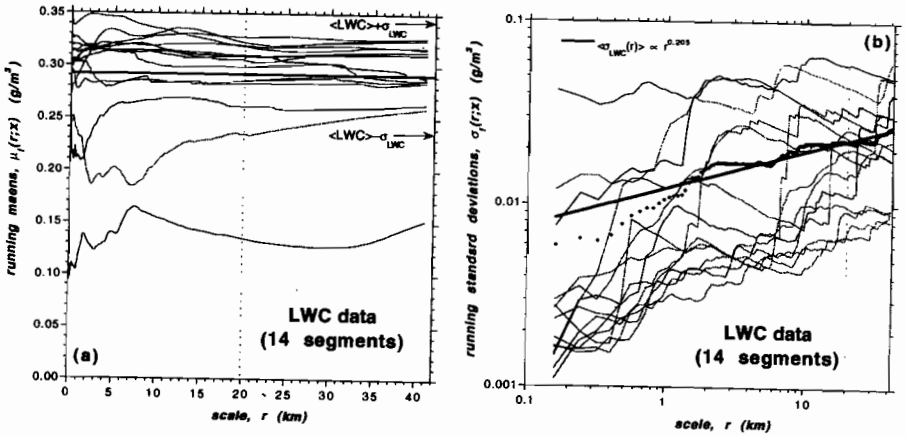


Figure 8: *Running Averages for 1D LWC Data.* (*a*) $\mu(r;x)$ vs. r for $0 \le r \le 40$ km and 15 different locations separated by 40 km or more. (*b*) Same as (*a*) but for $\sigma(r;x)$.

[1]We are in no position to claim that 0.29 g/cm^3 is climatologically relevant, not even for marine Sc in summer off of southern California (where FIRE was conducted in June-July 1987). Moreover, even for a given cloud-type, location, and season, we have 5 means (also true of characteristics described below).

4.4 Multifractal Analysis 2, Singular Measures

Figure 9*a* shows the spatial/ensemble-average[1] $\log_2\langle\varepsilon(r;x)^q\rangle$ vs. $\log_2 r$ for integer-valued $q = 1,\ldots,5$ and $\log_2(r/\eta) = 0,\ldots,[\min_{i=1,\ldots,5}\log_2(L_i/\eta)]$ using the LWC data in Eq. (2) with $\eta \approx 4\ell$ (20 m) and then Eq. (15). For $r \geq \eta$, we have power laws for all q's; adopting the notation of Schertzer and Lovejoy (1987), we posit:

$$\langle\varepsilon(r;x)^q\rangle \sim r^{-K(q)} \tag{20}$$

for the qth-order moments of the coarse-grained measures. In Fig. 9*b* we show $K(q)$ versus q for $0 \leq q \leq 5$.

An often used alternative to the d-dimensional *coarse-grained* measure $\varepsilon(r;x)$ is the *total* measure in the interval $[x,x+r)$ or domain $[x,x+r)\otimes[x,x+r)$:

$$p(r;x) = r^d\varepsilon(r;x). \tag{21a}$$

In this case, sums —not averages— of the qth power of $p(r;x)$ over the $(L_2/r)^d$ disjoint intervals are used:

$$Z_r(q) = \sum_x\langle p(r;x)^q\rangle \sim r^{\tau(q)}, \tag{21b}$$

where $\langle\cdot\rangle$ now designates an (optional) ensemble-average over different realizations of ε. By separating spatial and ensemble averages in (20) $(L_2/r)^d\langle\varepsilon(r;x)^q\rangle = r^{-dq}\sum_x\langle p(r;x)^q\rangle$. Hence, from Eqs. (20) and (21*b*), we find

$$\tau(q) = (q-1)d - K(q). \tag{22}$$

Methods for estimating $\tau(q)$ using continuous wavelet transforms have been developed by Arnéodo *et al.* (1988).

Two $K(q)$ values are predetermined. Normalization requires $K(0) = 0$ ($\tau(0) = -d$). Only at $q = 1$ can we permute the spatial averaging inside the interval $[x,x+r)$ and the spatial average over the various x's. In the convention where x is sampled every r, there is no difference in the outcome as r is varied: $K(1) = 0$ ($\tau(1) = 0$).

We note the convexity of $K(q)$ in Fig. 9*b*, the associated $\tau(q)$ being concave. This remarkable property is traceable to the same probabilistic cause as for $\zeta(q)$'s concavity. Here again this can be used to define the non-increasing hierarchy of exponents:

$$D(q) = \frac{\tau(q)}{q-1} = d - \frac{K(q)}{q-1}, \tag{23}$$

the generalized dimensions introduced originally by Grassberger (1983) and Henchel and Procaccia (1983) for the characterization of strange attractors in deterministic chaos theory.[2] Some dimensions are noteworthy; in standard terminology, we have:

[1] It is customary to resample x every r (from 0 to L_2-r) in the spatial part of the averaging procedure; this guarantees that no data is used more than once for a given statistic.

[2] In that context, $p(r;x)$ is the number of points sampled in the phase space of the dynamical system.

- $D(0^+) = d+K(0^+) = \tau(0^+)$, dimension of the support;[1]
- $D(1) = d-K'(1) = \tau'(1)$, information dimension;
- $D(2) = d-K(2) = \tau(2)$, correlation dimension.

At $q = 2$, the scaling of $\langle\varepsilon(x+r)\varepsilon(x)\rangle$ can be related to that of $\langle\varepsilon(r;x)^2\rangle \sim r^{-K(2)}$, at least for multiplicative cascade models (see Appendix, §A.3.2). From the power-law translation of the Fourier duality in Eq. (9*b*), we have

$$\beta_\varepsilon = 1-K(2) = 1-[d-\tau(2)] = (1-d)+D(2). \tag{24}$$

A fourth q-value was attracted considerable attention:

- $D(q_D) = 0$ or $\tau(q_D) = 0$ or $K(q_D) = (q_D-1)d$ define the "critical" moment q_D.

For $q \geq q_D$, the moments of $\varepsilon(r;x)$ are divergent (Mandelbrot 1974, Kahane and Peryière 1976, Schertzer and Lovejoy 1987, Gupta and Waymire 1993, and others).

Data like ours, yielding a nonlinear $K(q)$ hence a non-constant $D(q)$, are "multifractal" in the sense of Halsey *et al.* (1986), Meneveau and Sreenivasan (1987*a,b*), Schertzer and Lovejoy (1987), Evertsz and Mandelbrot (1992), and others. If $D(q) \equiv$ *constant*, $K(q)$ and $\tau(q)$ are proportional to $q-1$, and the data are said to be "monofractal."

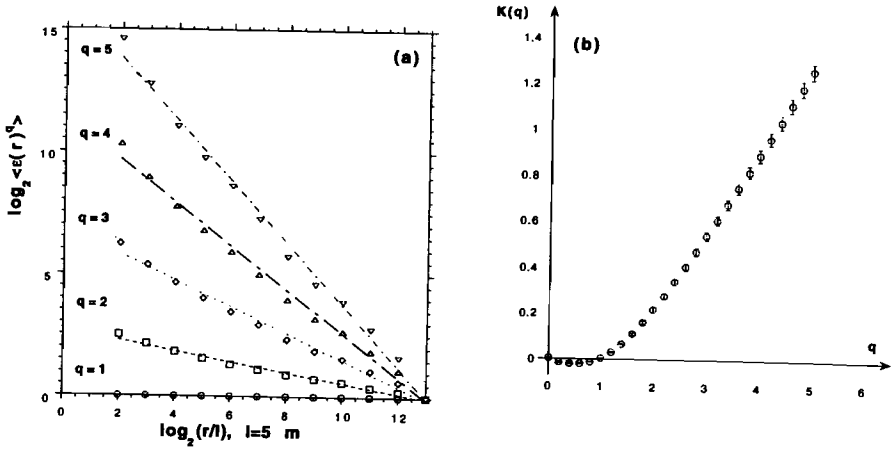


Figure 9: *Singular Measures for 1D LWC Data.* (*a*) Scaling for $q = 1(1)5$. (*b*) $K(q)$ exponents.

4.5 Bi-Fractal Analysis

Is it strictly necessary to have two kinds of multifractal analyses? In other words: *Is there a general connection between $\zeta(q)$ and $K(q)$*? This is an open question discussed at length in the specialized literature (Schertzer and Lovejoy 1987, Meneveau and Sreenivasan 1991, Sreenivasan 1991, Frisch 1991, Davis *et al.* 1993, Vainshtein *et al.* 1994), the consensus being that at least one extra exponent is needed to go from the

[1] $q = 0^+$ means that $\varepsilon(r;x) > 0$ gives 1 and $\varepsilon(r;x) = 0$ gives 0. Note that $K(0^+) \leq 0$, due to convexity.

currently fashionable $K(q)$-based approach to the more traditional one based on structure functions and $\zeta(q)$. Let this exponent be

$$H_1 = H(1) = \zeta(1); \tag{25a}$$

Bertozzi and Chhabra (1994) show that $\kappa = 1-H_1$ is the "cancellation" exponent recently introduced by Ott *et al.* (1992) to measure the rate at which $f(x)$ goes from an up-trend to a down-trend. We prefer to view the (mean) Hölder exponent H_1 as an index of nonstationarity: the degree of nonstationarity increases with H_1, ranging from 0 (stationarity) to 1 (differentiability).[1]

Another question of practical importance is: *How many q's do we actually need in any given application?* It is impossible to answer this question on general grounds. However, we consider it important in any application to get at least a first-order estimate of the degree of intermittency in the system. Using the small-scale absolute gradients is a convenient way of doing this and a 1st-order index of intermittency is the information codimension

$$C_1 = d-D(1) = d-\tau'(1) = K'(1). \tag{25b}$$

At $C_1 = 0$, there is no intermittency. At $C_1 = d$, intermittency is maximal, corresponding to a situation where all the measure is concentrated onto a finite number of points (cf. Dirac measures in the Appendix).

Relations of the type

$$\zeta(q) = q/a-K(q/b) \tag{26}$$

have been proposed where a and b are constants. For instance, $a = b = 3$ for turbulent signals (e.g., Sreenivasan 1991) and $a \le b$ for a model of Schertzer and Lovejoy's (1987) described in the Appendix. The generality of this $\zeta(q) \leftrightarrow K(q)$ connection remains an open question.[2] If however Eq. (26) is either true in general, accepted as an approximation for low enough q's, or used as a definition of $K(\cdot)$ after setting b, then one can derive both H_1 and C_1 from structure functions alone, without resorting to measures based on gradients. Indeed, Eqs. (25–26) yield

$$\begin{cases} H_1 = \zeta(1) = 1/a-K(1/b) \\ C_1 = K'(1) = b/a-b\zeta'(b) = \zeta(b)-b\zeta'(b) \end{cases} \overset{b=1}{\rightarrow} \begin{cases} H_1 = \zeta(1) = 1/a \\ C_1 = K'(1) = \zeta(1)-\zeta'(1) \end{cases} \tag{27}$$

Notice that if $\zeta(q) \propto q$, then $K(q) \equiv 0$, hence $C_1 = 0$; so C_1 is a measure of the curvature in $\zeta(q)$. As an example, we obtain $C_1 \approx 0.06$ in this manner for the 2D Landsat data using $b = 1$ for simplicity (cf. Fig. 7).

"Bi-fractal" analysis is the minimal form of multifractal analysis based solely on H_1 and C_1. [Davis and Marshak (1996) discuss, compare, and relate other choices used in the literature.] Figure 10 shows a schematic (H_1,C_1) plane for $d = 1$. In our experience with marine Sc, bi-fractal analysis has proven very useful, leading in particular to new

[1] In §5.1 we will show that processes with $\beta < 1$ are stationary; those with $1 < \beta = 2H(2) + 1 < 3$ are nonstationary with stationary increments; those with $\beta > 3$ are almost everywhere differentiable and have nonstationary increments. Since $H(2) \le H(1) = H_1$, we have $0 < H_1 < 1$.

[2] A methodology for empirical verification is suggested by Davis *et al.* (1993).

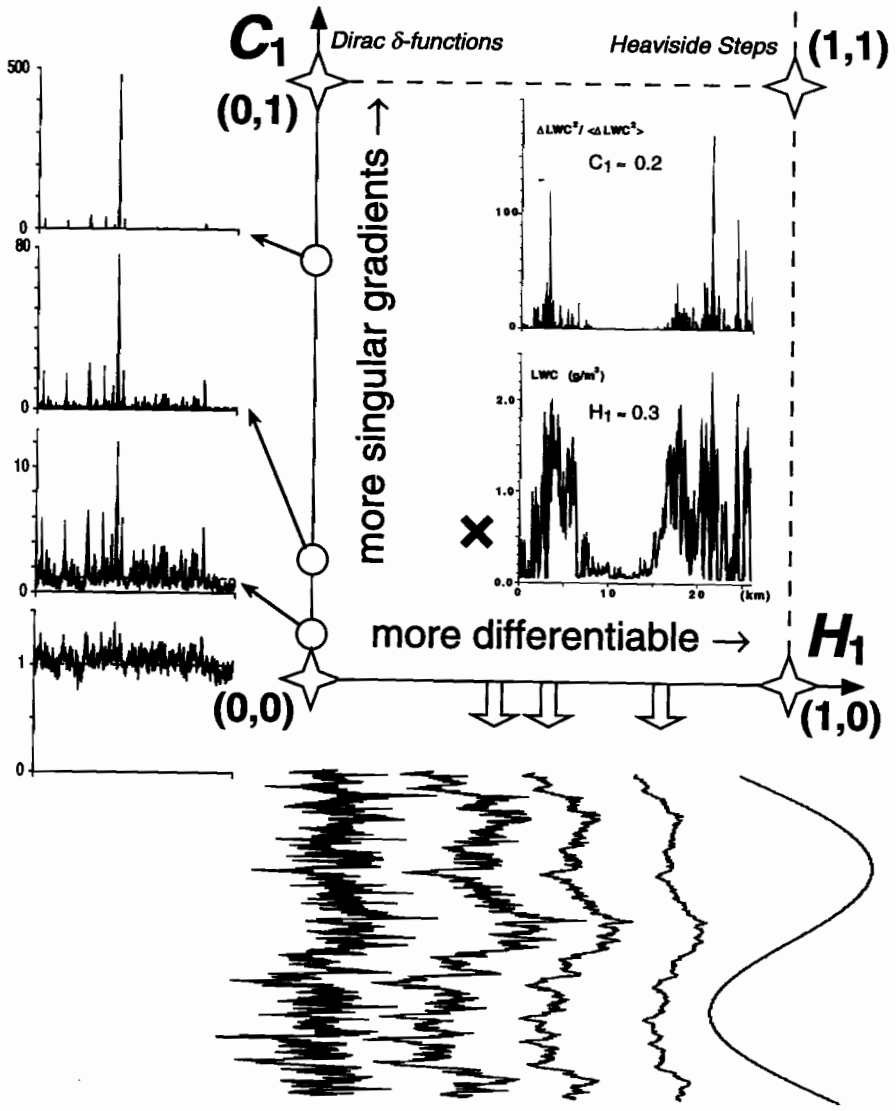


Figure 10: *The Bi-Fractal Plane.* This schematic shows how the vertical coordinate reflects the intermittency of the squared small-scale gradients (lower inset) whereas the horizontal coordinate characterizes the nonstationarity of the primary signal (upper inset). Synthetic data with increasing C_1 and H_1 grace the axes. We see that, contrary to spectral analysis, bi-fractal analysis distinguishes Brownian motion, $(H_1,C_1) = (1/2,0)$, from Heaviside steps, $(H_1,C_1) = (1,1)$, which are both nonstationary ($H_1 > 0$). The same is true for the stationary ($H_1 = 0$) gradients of these two theoretical cases: white noise has $(H_1,C_1) = (0,0)$, and a single spike, modeled by a δ-function, has $(H_1,C_1) = (0,1)$.

insight into cloud dynamical processes. Table 2 gives (H_1,C_1) coordinates for the data presented in section 2.

Table 2: *(H_1,C_1) values for FIRE Landsat radiances and LWC, ASTEX and SOCEX LWC, and ARM LW path data.*

Database (dataset)	H_1	C_1	r_{min} / r_{max}	Reference
FIRE Landsat radiances (6/30)	0.54	0.06	150 m / 10 km	(this publication)
FIRE LWC (6/30, 22:41Z)	0.29	0.14	20 m / 20 km	Marshak *et al.* (1996)
FIRE LWC (7/02, 02:23Z)	0.22	0.15	" "	" " "
FIRE LWC (7/14, 23:09Z)	0.34	0.03	" "	" " "
FIRE LWC (7/16, 17:17Z)	0.31	0.08	" "	" " "
FIRE LWC (7/16, 18:19Z)	0.34	0.07	" "	" " "
FIRE LWC 1987 ensemble	0.28	0.10	" "	" " "
ASTEX LWC 1992 ensemble	0.29	0.08	60 m / 60 km	Davis *et al.* (1994*a*)
SOCEX LWC 1994 ensemble	0.28	0.09	5 m / 5 km	(in preparation)
ARM LW path ensemble	0.37	0.08	1 min / 8 hr	Wiscombe *et al.* (1994)

The scatter, most notable in the intermittency index, reflects the diversity in Figs. 2*a–e* and argues for a non-ergodic model for this data. FIRE was conducted off the coast of Southern California, near San Diego. We have added to Table 2 (H_1,C_1) entries for ensemble averages from analyses of LWC in marine Sc for two other field programs: the Atlantic Stratocumulus Transition EXperiment (ASTEX) that focused on a more complex situation with transitions to cumulus regimes, and the Southern Ocean Coupled EXperiment (SOCEX) that was conducted off the coast of Australia. The proximity of the ensemble average (H_1,C_1) points for the three different local climates argues for a degree of universality in the dynamics that determine the internal structure of marine Sc. The H_1's for LWC cluster near 1/3, the value that characterizes turbulent fields: velocity (Kolmogorov 1941), temperature, and passive admixtures (Obukhov 1949, Corrsin 1951). The C_1's for LWC cluster near 0.1, precisely in the range observed for turbulent signals.[1,2] Thus, although H_2O (in all its phases) is far from being a dynamically passive constituent of the atmosphere, it can be perceived as advected by the turbulent wind field to a first approximation.

Finally, we have appended to the above LWC statistics an ensemble-average (H_1,C_1) for liquid water path (column integrated LWC) measured at the ARM site in Lamont (Ok.) for arbitrary cloud cover, as opposed to Sc only. We find roughly the same C_1 but a somewhat larger H_1 than for the three LWC observations. This is not surprising: the vertical integration that relates LWC to LW path will produce a smoother signal. LW path is more directly relevant than LWC to the radiative properties of the cloud, as

[1] C_1 values can be obtained from "intermittency parameters," $K(2)$, gleaned in the literature. To go from the characterization at $q = 2$ to $q = 1$, two extreme hypotheses are log-normality, $K(q) = C_1q(q-1)$, and monoscaling, $K(q) = C_1(q-1)$. Allowing for this uncertainty ($C_1/K(2) \approx 1$–2), soC_1 falls in 0.2–0.3.

[2] In turbulence studies, the dissipation rate field is obtained by squaring the velocity gradients at the Kolmogorov scale, leading to $K_2(q)$, rather than taking their absolute values, leading to $K_1(q)$, a difference in methodology easily accounted for: Lavallée *et al.* (1993) show that $K_1(q) = K_2(q/2)-qK_2(1/2)$ which puts $C_1 = K_1'(1)$ in the range 0.07–0.15.

observed in the Landsat image. The H_1 for marine Sc radiance fields in visible channels is even smoother than for LW path because the escaping radiation fields are highly scattered. Davis *et al.* (1996*b*) show that multiple scattering leads to a non-trivial physical smoothing over scales ≈200–300 m, and Marshak *et al.* (1995*b*) use multifractal analysis to show that this smoothing affects large and small jumps in the horizontal distribution of LW path differently.

5. Semi-Empirical Criteria (Statistical Interpretation of Scaling Regimes and Exponents)

In this section, we state or establish some results, typically inequalities between exponents, that enable us to classify data as stationary or not, ergodic or not, intermittent or not, according to the scaling behavior of various statistics. The same data can have conflicting attributes (e.g., stationary and nonstationary) as long as they refer to different ranges of scales. From a statistical standpoint, it will become apparent that stationarity, ergodicity, and intermittency are just different facets of the basic issue of data analysis: *What properties should we determine from our data?* and *How accurately can we estimate them from our finite sample?* From a physical standpoint, stationarity and intermittency are clearly more fundamental concepts than ergodicity.

5.1 Criterion for Stationarity (A Necessary Condition for Ergodicity)

In the Appendix, we describe procedures for synthesizing a number of scale-invariant models for stochastic processes and discuss their properties, among these "stationarity." This body of theoretical knowledge is important for a variety of reasons. First, models enable validation of analysis software —"instrumental calibration" in our laboratory analogy— as well as sensitivity studies (how does the output depend on the amount and properties of the data being processed?). Second, models can be used in applications (e.g., cloud radiation studies). Last but not least, models generally have well-understood properties. Strictly speaking, stationarity is a property that can only be assigned to a model because the question is: *Are statistical quantities, as defined by ensemble averaging (i.e., over probability- or f-space), invariant under translation in x?* This is not easy to address with data because we always operate with finite amounts of data; furthermore we rely heavily on spatial averages to estimate statistics in the first place, so at least some x-dependence is operationally erased.

Another theoretical question, more directly relevant to data analysis, is that of "ergodicity:" *Do spatial averages of increasing length for a single realization converge to ensemble averages (over all possible realizations)?* In data analysis we generally make implicit ergodicity assumptions before computing spatial averages, i.e., we expect them to converge to something meaningful. We also assume implicitly that this definite number we are seeking does not depend on when we start computing it (stationarity). It is therefore important to have guidelines as to what quantities are statistically well-defined, not just computable by some given algorithm.

Without exception, the scale-invariant models in the Appendix with

$$\beta < 1 \qquad (28a)$$

are stationary in the "broad" sense where $G(r) = \langle [f(x+r)-\langle f\rangle][f(x)-\langle f\rangle]\rangle$ depends only on r; we recall that, in this theoretical context, the 2nd-order autocorrelation function is obtained by averaging over all possible (or at least many) f's, holding x constant. Davis *et al.* (1994*b*, 1996*a*) give more general arguments for drawing the line between stationary and nonstationary behavior at $\beta = 1$ for scale-invariant processes.

If we only have

$$\beta < 3 \qquad (28b)$$

then the model has (broad-sense) stationary increments: $\langle [f(x+r)f(x)\rangle = G(r)+\langle f\rangle^2$, and even $\langle f(x)\rangle$, may or may not depend on x but the structure function of order $q = 2$, namely $\langle [f(x+r)-f(x)]^2\rangle$, does not. We propose to use the criteria in Eqs. (28*a*,*b*) for real world data-streams as well as for theoretical models. For data, $\beta > 1$ means that spatial estimates of $\langle f(x+r)f(x)\rangle$ are likely to vary from one realization (or portion of data) to the next. For data, $\beta < 3$ means that spatial estimates of $\langle [f(x+r)-f(x)]^2\rangle$ are likely to be robust (i.e., invariant under addition of new data).

In this scheme, the FIRE LWC data ($\beta \approx 1.4$) in Figs. 2*a–e* is nonstationary with stationary increments for scales from ≈20 m to the integral scale ≈20–40 km. For larger scales, we have no spectral information in Fig. 4*c* but the leveling-off of the structure functions in Fig. 6*a* confirms this estimate of the integral scale. This is symptomatic of stationarity (increments cease to grow). The FIRE Landsat data ($\beta \approx 2.0$) in Fig. 3 is also nonstationary with stationary increments from the radiative smoothing scale ≈200 m to ≈10 km. Were it not for saturation at maximal gray level 255, spectral flattening (transition to stationarity) would occur closer to the integral scale for the cloud LW path fluctuations, itself likely to be around that of LWC.

In contrast, the $\varepsilon(\eta;x)$ fields in Figs. 2*a'–e'* ($\beta \approx 0.7$) are stationary in spite of their intense spikiness. This may seem surprising since local means, $\varepsilon(r;x)$ with $r \gg \eta$, will fluctuate wildly, depending on the strength and number of spikes that fall in the interval $[x,x+r)$. The strong variability of these fields therefore contrasts with the conventional wisdom about stationarity, essentially that (in the usage of time-series analysis)

'temporal statistics do depend little on when they are gathered.' (*)

The occurrence of spikes of course perturbs strongly the local statistics and therefore violates this operational definition of stationarity. We prefer to think of property (*) as a consequence of ergodicity (which is more restrictive than stationarity): if, in general, running temporal averages converge reasonably fast to their ensemble counterparts as the sample size increases, then clearly we need not worry about where we start cumulating. The two unstated assumptions in effect here are:

1) 1-point statistics are Gaussian-type, i.e., that only relatively small deviations from mean or modal values are anticipated;
2) 2-point correlations are of short range, i.e., the integral scale is relatively small.

Samples of relatively small length therefore provide enough data to obtain accurate estimates of 1-point and 2-point statistics of all orders.

In our outlook, stationarity should have no bearing at all on the intensity of the fluctuations; it should however have a strong impact on their "rapidity" (since the smaller the spectral exponent the more variance in the small scales). In short, we propose to drop the above assumption #1. Even in strongly intermittent cascade processes $\varepsilon(x)$ with $\langle\varepsilon\rangle = 1$, spikes (where $\varepsilon(x) \gg 1$) do not prevent the signal from

'having a well-defined mode,' (†)

(which is naturally $\ll 1$). Property (†) is, in our opinion, a better description of a stationary time-series, whether Gaussian/ergodic or not.[1]

5.2 The Onset of Sampling Problems (Trivial Ergodicity Violation by Finite Datasets)

In the above considerations, only 2nd-order statistics are used, and we assume that the data is ergodic for that type of statistic (dominated by relatively frequent events). Going to ever higher-order moments, extreme events gain more and more weight. In estimators based on a finite amount of data, a single event will eventually dominate. This can be sensed in scaling analyses when the exponent functions, $\zeta(q)$ or $K(q)$, become linear in q. This is a gradual transition but one can nevertheless define q_S such that

$$\zeta'(q) \text{ or } K'(q) \approx \textit{constant} \text{ for } q > q_S.$$

Schertzer and Lovejoy (1992) derive an expression for q_S in the frame of singular measures. Now some singular cascade models for $\varepsilon(r;x)$ (e.g., "p-models" discussed in the Appendix) are immune to sampling problems yet their $K(q)$'s become asymptotically linear in q, simply because there is a well-defined maximal event present in every realization. In practice, one can easily test the hypothesis of obtaining roughly the same $\max_x\{\varepsilon(r;x)\}$ for every realization. So, in principle, there is no risk of misinterpreting the observation of linear trends in $K(q)$, at least for models.

Using the FIRE LWC data, Figs. 6*b* for $\zeta(q)$ and 9*b* for $K(q)$ are quasi-linear for $q \gtrsim 3$, which is our estimate for q_S. The $\zeta(q)$ in Fig. 7*b* for the FIRE radiance data is quite linear for small q's (see below) but does not show an asymptote for large q.

5.3 Criterion for Intermittency (Non-Trivial Ergodicity Violation by Non-Gaussian Processes)

Suppose we want to model data (or the geophysical field it samples) with the simplest possible scale-invariant stochastic process. A question naturally arises: *'Is a non-Gaussian multifractal model strictly necessary or is a simpler monofractal one good enough?'* Focusing on qth-order structure functions in Eq. (16), we find mono- (or "simple") scaling —namely, $\zeta(q) = \zeta(1)q$— as soon as the trivial dimensional reasoning, that is

$$\langle |f(x+r)-f(x)|^q\rangle \approx \langle |f(x+r)-f(x)|\rangle^q, \qquad (29a)$$

[1] A random sample is, by definition, dominated by modal values and will generally not give good estimates of high-order statistics, even of the mean if the distribution is sufficiently broad or skewed.

makes accurate predictions. Similarly, for singular measures, we would have

$$\langle \varepsilon(r;x)^q \rangle \approx \langle \varepsilon(r;x) \rangle^q, \tag{29b}$$

hence $K(q) \approx K(1)q \equiv 0$. We have argued above that, for all practical purposes (i.e., a finite amount of data is available), Gaussian-type increments $f(x+r)-f(x)$ or $\varepsilon(x)$ fields will indeed yield good estimates of relatively high-order moments that obey (29*a*,*b*). This type of data can be deemed "ergodic" and is necessarily non-intermittent in the sense that $K(q)$ is vanishingly small. It is therefore important to give a quantitative meaning to the "$\approx$" symbols in Eqs. (29*a*,*b*)?

In theory, fractal and multifractal properties require a small-scale limit ($r \to 0$) to become mathematically precise. In practice, multifractal analyses are performed with finite amounts of data with a finite range of scales. There will almost[1] always be trivial ergodicity violations based on extreme events, as described briefly in the previous subsection. In turn, these events lead to a small but finite degree of multiscaling. *Can we distinguish between this spurious multifractality and its "real" counterpart (i.e., that is likely to be robust under addition of new data)?*

This exercise in ergodicity verification normally requires either obtaining more data or subdividing the available data and re-doing the analyses in order to monitor the effect of sample-size. Aurell *et al.* (1992), Marshak *et al.* (1994), Eneva (1994), and others have explored this sampling issue analytically or numerically with specific models that yield $K(q) \equiv 0$ in the small-scale limit but $K(q) \not\equiv 0$ for a finite range of scales. Grivet-Talocia (1995) and others have investigated finite-size effects that cause stationary scaling processes with $\beta < 1$ to have small but finite $\zeta(1)$ and $\zeta(2)$, in contradiction with Eq. (18), let alone $\zeta(q) \equiv 0$. A general (model-independent) first-order answer to the above question is now derived from the hypotheses in Eqs. (27*a*,*b*).

Let $\xi(r;x) = |f(x+r)-f(x)|$ or $\varepsilon(r;x)$, depending on the multifractal approach of interest. Equations (29*a*,*b*) become

$$\langle \xi(r;x)^q \rangle^{1/q} \approx \langle \xi(r;x)^{q_{\text{ref}}} \rangle^{1/q_{\text{ref}}}. \tag{30}$$

For specificity, we can take $q_{\text{ref}} = 1$ but in some applications $q_{\text{ref}} = 2$ would be a natural choice. In the following, it is assumed that some range $[q_{\min}, q_{\max}]$ is explored for a number of scales, ranging from $r_{\min}$ to $r_{\max}$. To make the "$\approx$" sign in Eqs. (29–30) quantitatively meaningful, we just need to make sure that there is minimal information in the prefactors on q-dependence. There often is (cf. Figs. 6*a* and 7*a*), but of a trivial kind, just reflecting a poor choice of units. It is easy to select physical units for ξ that remove this trivial dependence for singular measures: by normalizing $\varepsilon(r;x)$ so that the r-independent statistic $\langle \varepsilon(r;x) \rangle$ is unity (cf. Fig. 9*a*). For structure functions, there is no r-independent case but we can always choose units that make $\langle |f(x+r_{\max})-f(x)|^{q_{\text{ref}}} \rangle = 1$; then, if there is little statistical information on q in the prefactors, the same should be approximately true in general: $\langle |f(x+r_{\max})-f(x)|^q \rangle \approx 1$. Visual inspection of log-log plots of $\langle \xi(r;x)^q \rangle$ vs. r is enough to show that this is possible: the intercepts of regression lines should be approximately linear in q.

[1] We exclude models that are cunningly ergodic in the sense of one or the other of the multiscaling statistics (cf. "p-model" in the Appendix for singularity analysis).

Using the definitions in Eqs. (16) or (20), we can rewrite the constraint for monoscaling in Eq. (30) as $(r/r_{\max})^{A(q)} \approx (r/r_{\max})^{A(q_{\text{ref}})}$ where

$$A(q) = \begin{cases} \zeta(q)/q = H(q) \\ -K(q)/q = \tau(q)/q + (1/q-1)d = (1/q-1)[d-D(q)] \end{cases}. \tag{31}$$

In the (lower) singular measures case, $A(q)$ is a new non-increasing function.

Now we must decide at what degree of discrepancy in Eq. (30) we are compelled to take multiscaling into account. There is no universal answer for this at present; in particular, the number of samples will be a factor. However, there is a general consensus that multifractals are a framework for modeling "strong" variability. It seems reasonable to require that the ratio between both sides of Eq. (30) be less than, say, one order-of-magnitude or more at the opposite end of the range of scales from where their quantitative agreement can be imposed arbitrarily, simply by using the proper units. We did this at $r = r_{\max}$, so we will now focus on $r = r_{\min}$. Conversely, we can require that *bone fide* multiscaling data obey

$$\langle \xi(r;x)^q \rangle^{1/q} / \langle \xi(r;x)^{q_{\text{ref}}} \rangle^{1/q_{\text{ref}}} \big|_{r=r_{\min}} \approx (r_{\min}/r_{\max})^{A(q)} / (r_{\min}/r_{\max})^{A(q_{\text{ref}})} \gtrsim B \tag{32}$$

where B is an arbitrary but relatively "big" number, like 10 or 100 (depending on the application, amount of data, etc.). Taking logs in Eq. (32) yields

$$\log\left(\frac{r_{\max}}{r_{\min}}\right) A(q) \Big|_{q}^{q_{\text{ref}}} \gtrsim \log B. \tag{33}$$

The chance of passing this test is clearly greater if (*i*) the range is scales $r_{\max}/r_{\min}$ is increased, or (*ii*) the range of q's is increased.

We now restrict our attention to singular measures with which we characterize intermittency routinely. The criterion in Eq. (33) can be translated into another for C_1 if we make an assumption on $A(q)$, hence $K(q)$. For small enough values of q, the log-normal model, $K(q)/C_1 = q^2-q$ (cf. Appendix), is a reasonable fit to our generally parabola-shaped $K(q)$ curve (e.g., Fig. 9*b*); we can therefore take $A(q)/C_1 = 1-q$. Combining this simple expression with the inequality in (33) and maximizing the range of q (i.e., $q = q_{\max}$, $q_{\text{ref}} = q_{\min}$), this $A(q)$ function yields a simple —erring somewhat on the conservative side— criterion where we can isolate C_1 as the only data-dependent quantity, the others being essentially instrumental: the data is "truly" multifractal only if

$$C_1 \gtrsim \frac{\log B / \log(r_{\max}/r_{\min})}{q_{\max}-q_{\min}}. \tag{34}$$

We have $r_{\max}/r_{\min} \approx 10^3$ for most of the entries in Table 2. The denominator in the r.h.s. is 5; taking $B = 10$, we require $C_1 \geq 0.07$ to qualify the data as truly multiscaling. One case (FIRE LWC on 7/14) fails the multifractality test; being so exceptionally smooth, this is not a surprise. The others cases pass, although one just border-line.

We noted already that the $\zeta(q)$ results in Fig. 7*b* for the FIRE radiance fields are almost linear in q for small enough q (say, 0 to 2.5). Can we argue for multifractality with this data? In this case, we have $r_{\max}/r_{\min} \approx 10^2$ (this is a somewhat conservative

estimate because of the large-scale effect of radiometry saturation). Values of $A(q)$ for $q_{ref} = 1$ and $q_{max} = 5$ are $A(1) = \zeta(1) = H_1 \approx 0.5$ and $A(5) = \zeta(5)/5 \approx 2.0/5$. Using base 10 logarithms, the l.h.s. of Eq. (33) is therefore $\approx 2^+ \times (0.5\text{-}0.4) = 0.2^+$, much less than the r.h.s. for $B = 10$. This argues for a monofractal model for this data. Surprising result since we have good evidence of multifractality in the associated LWC fields which are representative of the fluctuations of the extinction (photon scattering probability per unit of length) in the cloud. Numerical simulations by Marshak *et al.* (1995*b*) explain this paradox: multiple scattering processes smooth the large jumps in extinction more effectively than the small ones; the latter determine H_1 and the former, the higher-order moments of the increments.

6. THEORETICAL CONSIDERATIONS (Geophysical Data Analysis as a Problem in Statistical Physics)

Experimentation is generally conducted in the laboratory with a theoretical model in mind, a hypothesis to test. At the level of generality that we have adopted in our survey of multifractal data analysis techniques, the most pressing theoretical questions are: *'Why is scaling almost universally observed in geophysical signals and fields?'*; and *'What can we learn about the underlying physical processes from the scaling properties?'*.

6.1 Thermodynamical Interpretation of Multifractal Quantities

In §3.3 we likened the statistical parameter q with a standard one in experimental work, namely ambient temperature. In the same vein, there is an increasingly popular interpretation of all multifractal exponents as thermodynamical quantities, first explored by Feigenbaum (1987) and recently surveyed by Muzy *et al.* (1994). In particular, diverging moments (important for modeling ergodicity/sampling problems) are perceived as 1st- or 2nd-order phases transitions, their signature being a discontinuity in the 1st- or 2nd-order derivatives of the "equilibrium curves" $\zeta(q)$ or $K(q)$; see Schertzer and Lovejoy (1992) for a discussion in the frame of singular measures.

There are solid physical reasons for exploiting this formal analogy. Indeed, the current rationale for using scaling analysis is that geophysical systems by nature have very many interacting degrees-of-freedom. We can think of the number of computational cells required to solve the coupled PDEs for Navier-Stokes equations, generally with ancillary constraints, for very high Reynolds numbers. In this respect, we imagine the generally large, complex geophysical system under consideration as a thermodynamical one: it has many allowable configurations in the sense that all the microscopic variables can each take on a number of values. Two classic examples are: positions, linear- and angular momenta of molecules in a gas; spin values of atoms or domains in a magnetic material. This defines a vast probability space, impossible to describe in any kind of detail. However, the macroscopically observable quantities are few (akin to temperature, free energy, entropy, magnetic induction, etc.) and are defined by ensemble-averages over all possible microscopic configurations. These "observables" generally depend little on the detailed dynamics of the system and thus define universality classes (e.g., real gases, Ising models). The counterpart of the thermodynamical limit in statistical physics (i.e., very many interacting particles) here is the limit of a huge computational grid (hence a large range of scales) and we again expect some kind of universal behavior to arise.

6.2 Information Created by Breaking the Scaling Symmetry

Scaling can be viewed as a symmetry (or invariance) obeyed by the macroscopic geophysical system probed during data collection. In this case, we are dealing with an invariance under change of scale. There are other possible symmetries: "stationarity" (§5.1), invariance under translation; and "isotropy" (§3.1), invariance under rotation in 2D, time reversal or parity ($x \rightarrow -x$) in 1D. Of course these symmetries are all of a statistical nature since exactly translationally and rotationally symmetric fields are constant. Generally speaking, the more symmetric the system, the less information is required to describe it. In our case, similar statistical properties for a whole range of scales are described by a single exponent (and a prefactor). It can be argued that information about any system can be gained only by breaking its symmetry. For instance, to measure the circumference of a circle, a mark must be placed somewhere (thus breaking its rotational symmetry). The first and last points in a time-series are special with respect to translations (a degree of nonstationarity is therefore introduced).

Our experimentation with cloud data has confirmed this. In discussing Table 2, we highlighted the similarity of $q = 1$ multifractal exponents for the same type of cloud (marine Sc) at three different locales (FIRE data from the N–E Pacific, ASTEX data from the Mid-Atlantic, SOCEX data from the S–W Pacific). So a "bi-fractal" characterization tells us nothing about the local climatology; it does tell us however something about the universality of the thermo-hydrodynamical processes that shape the internal structure of marine Sc layers. In contrast, the scaling range, which is defined by scale-breaks at either end, does vary; it appears to be roughly in proportion with the thickness of the boundary layer (Davis *et al.* 1996*a*, Marshak *et al.* 1996).

From the cloud radiative perspective, the most interesting scale-break is the one at $\approx$200 m in Figs. 5 and 7*a* relating to the reflected radiance fields of marine Sc. Indeed, there is no counterpart of this statistically robust feature in the LWC data measured inside the same type of cloud. Davis *et al.* (1996*b*) survey the literature on this scale-break in the energy spectrum and they describe the radiative smoothing mechanism that produces it, via multiple scattering. Marshak *et al.* (1995*b*) use multifractal methods to investigate its implications for cloud remote sensing. In both studies, the phenomenology of the Landsat scale-break is based on a numerical Green's function analysis of horizontal photon transport, uncovering its dependence on geometrical cloud thickness and photon mean free path (corrected for the forward scattering): the scale-break occurs at the harmonic mean of these two fundamental scales.[1] Information about inherent cloud properties can therefore be extracted from the observation of the scale-break. Moreover,

[1] Interestingly, this finding is traceable to the effect of the non-illuminated cloud boundary at finite range on the photons' random walks that are modeled by *nonstationary* Brownian motion (cf. Appendix).

it is hoped that in the near future laser and low-light detector technologies will be combined to observe directly cloud radiative Green's functions. This would enable robust and inexpensive estimations of cloud thickness and density, both important quantities for balancing the Earth's radiative budget, hence forecasting climate change.

7. Summary

We have presented a conceptual model for geophysical data analysis based on laboratory work. The "sample" being probed is the data, generally collected in the field: either time-series, 1D transects, or 2D images. The laboratory "instruments" are computer codes that process this data and output 2- or more-point statistics, our focus being on scale-dependent statistical quantities that convey information about spatial correlations. The "readings" of these instruments are analyzed on log[*statistic*] vs. log[*scale*] plots, seeking straight lines that are the signature of scaling (power-law) regimes. A number of remarkable exponents (log-log slopes) are discussed and several criteria are presented: *When is the datastream stationary? ... When does it have stationary increments? ... Do we suffer from sampling ("ergodicity") problems? ... Is there enough intermittency to call for an inherently non-Gaussian "multifractal" model, as opposed to a simpler Gaussian one with "monofractal" statistics?* Finally, laboratory exercises are almost universally designed to validate or challenge the prevailing theory about the structure and dynamics of the sample. Experimental results can therefore prompt new theory. We show how this proves true in both cases for the theory of cloud-radiation interaction, an important pre-requisite in climate theory and the remote sensing of cloud properties. We therefore conclude that multifractal scale-by-scale analysis is a powerful —yet currently underexploited— tool in geophysical research, well-suited for connecting theory and measurements in a broad range of applications.

Appendix. Simulation and Calibration (Scale-Invariant Models: Stationary or Not, Intermittent or Not, Ergodic or Not)

An important, quasi-universal application of statistical data analysis, multifractal or other, is to constrain models that attempt to reproduce the data. So stochastic simulation tools are often developed in parallel with data analysis methods. This activity has both analytical and computational aspects. On the one hand, we need to write code that implement specific algorithms for generating random functions of one or two variables. On the other hand, we need to know, preferably in closed-form, the dependence of the statistical quantities of interest on the parameters of the model. In our experience, applications are two-fold:

- Stochastic cloud models have proved invaluable for investigating cloud radiation properties (e.g., reflectance, transmittance, and absorption of solar- and laser beams).
- Validation of data analysis algorithms: before applying them to real data, it is crucial to see how they respond to artificial data.

In the "laboratory" analogy for data analysis used in this paper, the "instrumentation" is a collection of computer programs that process the data-stream into statistics (large input, small output). Others will help fit nonlinear (e.g., power-law) models to the statistics in a more or less supervised manner. Just like real experimental procedures, these instruments must be calibrated with standard input, in our case, samples of data with known and controllable statistical properties.

In this Appendix, we present a comprehensive collection of theoretical models to perform this task. Simultaneously, the somewhat abstract concepts of stationarity and intermittency are made more palatable. On the one hand, we need both mono- and multiscaling nonstationary random *functions* with stationary increments to calibrate structure-function analysis. In section A.1, we recall how the simplest nonstationary processes are obtained from white noises in both ergodic and non-ergodic situations; in section A.2, we describe monoscaling fractional Brownian motions. On the other hand, we need stationary mono- and multiscaling random *measures* to calibrate singularity analysis procedures. In section A.3 such models are presented. In section A.4, we return to nonstationary functions to add two multiscaling models, both developed first for cloud studies, that complete the collection. In section A.5 finally, we summarize and display graphically connections between the different types of model.

Here again, we are dealing with a number of computer programs that synthesize data (small input, large output). This output is denoted $w(x_m)$, $f(x_m)$ or $\varepsilon(x_m)$ on a 1D grid of constant ℓ and size $M = L/\ell$:

$$x_m = m\ell, \qquad m = 0,\ldots,M-1. \tag{A0a}$$

Generalization to 2D grids,

$$(x_{m_1},y_{m_2}) = (m_1,m_2)\ell, \quad (m_1,m_2) = [0,M-1]\otimes[0,M-1], \tag{A0b}$$

are straightforward in most cases, at least if statistically isotropic models are acceptable. In one (more involved) case, the 2D construction is described in full detail (§A.3.5). Where convenient, we denote spatial dimensionality by

$$d = 1,2. \tag{A1}$$

Unless explicitly stated otherwise, we use units of length where

$$\ell = 1. \tag{A2}$$

For each model, we describe the generation algorithm in sufficient detail for direct coding; the main statistical properties are expressed as a function of their parameters, with either short derivations or references to the literature, and their stationarity, intermittency and/or ergodicity properties are discussed.

A.1 Nonstationarity in 1D, Running Sums of Stationary Processes

A.1.1 Gaussian White Noise and Brownian Motion

Brownian motion (Bm), a.k.a. the Wiener-Lévy process, is defined as the integral of white noise (w.n.), i.e., the running sum of a sequence of *uncorrelated* "steps" in the

forward or backward direction: the steps are denoted $w(x_m)$, $m = 1,\ldots,M$. These numbers are typically drawn from a Gaussian distribution but Bernoulli[1] and Laplace[2] distributions are also used. Throughout this Appendix, we denote by $N(\mu,\sigma)$ Gaussian[3] random variables of mean μ and variance σ^2 (standard deviation σ). The important requirement is finite variance to invoke the central limit theorem (CLT). For simplicity, we also require the vanishing mean to eliminate systematic drifts; so $w(x_m) = N(0,\sigma)$.

Bm is therefore

$$f(x_m) = f_0 + \sum_1^m w(x_j), \text{ no sum for } m = 0. \tag{A3}$$

If $w(x_m)$ is Gaussian —or, more generally (CLT), in the limit $m \gg 1$— $f(x_m)$ is Gaussian, and so are its increments $\Delta f(r;x_m) = f(x_{m+r}) - f(x_m) = \sum_{j=m}^{m+r} w(x_j)$, where $m = 0,\ldots M-r$ and $r = 0,\ldots,M$, because sums of Gaussians are Gaussian (with cumulated means and variances). It is easy to see that $\langle|\Delta f(r;x_m)|^q\rangle = \langle|\Delta f(r;x_0)|^q\rangle \propto \langle[f(x_r)-f_0]^2\rangle^{q/2} = [\sigma^2 r]^{q/2}$, where $\sigma^2 = \langle w(x_m)^2\rangle$ and[4] $q > -1$; thus $\zeta(2) = 1$ and, in general, $\zeta(q) = q/2$ from the definition of structure function exponents in Eq. (16). We now contrast the properties of w.n. and Bm:

- Gaussian w.n. is the prototypical *stationary process*: all its statistics are independent of the position x_m. Absolute 1-point moments are invariant, $\langle|w(x_m)|^q\rangle = \langle|w(x_1)|^q\rangle$ for $m = 1,\ldots,M$ and $q > -1$ as is the 2-point autocorrelation function

$$\langle w(x_{m+r})w(x_m)\rangle \equiv \sigma^2\delta(r), \tag{A4a}$$

  for $m = 1,\ldots,M-r$. W.n. is not only stationary but *ergodic*: spatial averages over a single but large enough realization converge to the above ensemble averages, largely because of the Gaussian nature of $w(x_m)$. In practice, very good estimates can be obtained from relatively small samples; see Fig. A1*a* for a case with $M = 1024$.
- Bm is the prototypical *nonstationary* process with stationary increments. Indeed, its 1-point statistics depend explicitly on m: we may have $\langle f(x_m)\rangle = \langle f_0\rangle$ (= 0 in the following) for all m but[5] $\langle f_m^2\rangle = \langle f_0^2\rangle + m\langle w_1^2\rangle$ for $m \geq 0$. The same is true of the 2-point autocorrelation:

$$\langle f(x_{m+r})f(x_m)\rangle \propto \langle f_0^2\rangle + |m+r|+m-r \tag{A4b}$$

  which follows from the identity $2f(x_{m+r})f(x_m) = f(x_{m+r})^2 + f(x_m)^2 - [f(x_{m+r})-f(x_m)]^2$. In

[1] A symmetric Bernoulli trial yields $w(x_m) = \pm s$ with equally probable signs, hence Bm on a grid (widely used in diffusion theory).

[2] In neutron or photon transport, Laplacian or two-sided exponential free paths (mean λ) are in order: $\pm\lambda\ln(\xi)$ with ξ uniformly distributed on (0,1) and equally probable signs.

[3] The Box-Muller transformation can be used to obtain zero-mean, unit-variance deviates; one way to do this is $N(0,1) = \cos(\pi\xi_1)(-2\ln\xi_2)^{1/2}$ where ξ_i $(i = 1,2)$ are computer-generated (pseudo-random) deviates distributed uniformly on (0,1). From there $N(\mu,\sigma) = \mu+\sigma N(0,1)$.

[4] Generally the r-dependent pdf, $\text{Prob}\{\Delta \leq \Delta f(r;x) < \Delta+d\Delta\}/d\Delta$, of the increment $\Delta f(r;x) = f(x+r)-f(x)$ $(r \geq 0)$ is non-vanishing at $\Delta = 0$; moments $\langle|\Delta f(r;x)|^q\rangle = \int\Delta^q \text{Prob}(d\Delta)$ are therefore divergent for $q \leq -1$.

[5] Setting $\langle f_0^2\rangle = 0$ enforces the nonstationarity (the $m = 0$ point is special). With $\langle f_0^2\rangle = M\langle w_1^2\rangle$, detrending and cyclically extending it, the resulting Bm is (cyclo-)stationary. However, we are interested primarily in scales $r \ll M$ for which this Bm is still nonstationary in the sense of the criterion in Eq. (28*a*): $\beta = 2 > 1$.

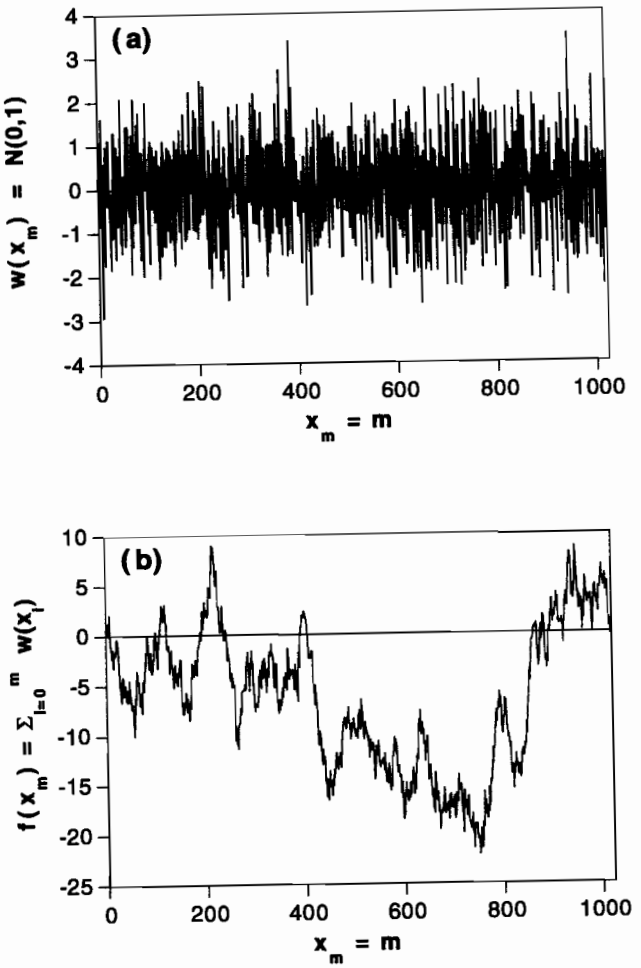


Figure A1: *Ergodicity, With and Without Stationarity.* (*a*) Uncorrelated (white) zero-mean unit-variance Gaussian noise; this data is not only stationary, but ergodic as well: spatial estimates of the low-order moments yield $\mu = 1.4\times10^{-4}$, $\sigma = 0.9962$ for $M = 1024$. (*b*) Running sum of the data in panel (*a*), i.e., Brownian motion; this data is nonstationary with stationary (and ergodic) increments.

contrast, 2-point statistics based on $\Delta f(r;x_m)$ depend only on r. Turning to spatial averages on a scale of r pixels ($1 \ll r \ll M$), 1-point means and variances will fluctuate wildly (non-ergodicity) but averages using $\Delta f(r;x_m)$ are well-behaved (ergodic increments) as long as Gaussian-type distributions are used. In summary, Bm is neither stationary nor ergodic *per se* but has stationary and ergodic increments. For an illustration, Fig. A1*b* shows the running sum of the data in Fig. A1*a*.

A.1.2 Lévy-Stable White Noise and Lévy Flights

White noise need not be Gaussian, hence so highly ergodic. Consider a power-law (e.g., Pareto) distribution where moments above some critical order are divergent; this means that their spatial estimates increase without bounds with sample size. Figure A2*a* illustrates this case with w.n. obeying a Cauchy law: moments of order $q \geq 1$ diverge. Only absolute moments of very small non-integer order can be properly estimated with $M \approx 10^3$. The counterpart of Bm for the Cauchian w.n. is graphed in Fig. 2*b*. This running sum of the data in Fig. 2*a* is a "Lévy-sable" process, often referred to as a Lévy "flight" (Mandelbrot 1983). Here incremental ergodicity is of course violated.

By definition, Lévy-sable variables are a four-parameter class of deviates $L(\alpha;a,b,c)$ obeying the rescaling equation:[1]

$$\sum_1^n L - na \stackrel{d}{=} n^{1/\alpha}(L-a), \tag{A5}$$

- α is the "Lévy index." For $\alpha \to 0^+$, the solution is $L = a$ (the degenerate distribution) and normally distributed random variables obey the rescaling in Eq. (A5) when $\alpha = 2$. In general, $0 < \alpha < 2$ is the critical order above which statistical moments diverge —variance included. For further details, we refer to Feller's (1971) treatise.
- a is a "centering" parameter, the counterpart of mean μ in Gaussian deviates.
- $b \in [-1,+1]$ controls skewness and has no equivalent in the Gaussian case.
- c is the amplitude parameter, like standard deviation σ in the Gaussian case.

We focus here on the most straightforward generalizations of the Gaussian case ($b = 0$); denoting these $L(\alpha;a,c)$, we have $N(\mu,\sigma) = \lim_{\alpha \to 2} L(\alpha;\mu,\sigma)$. The only case with a closed-form pdf is $\alpha = 1$ (Cauchy deviates); the sample in Figs. A2*a*,*b* also has $a = 0$.

A.1.3 Spectral and Multifractal Properties

We have already shown that Bm (Gaussian-type increments) is monoscaling with

$$H(q) \equiv 1/2, \; q > -1, \tag{A6a}$$

The spectral exponent in $E(k) \propto k^{-\beta}$ is related to the $q = 2$ case by $\beta = 2H(2)+1 = 2$.

The defining relation (A5) for non-skewed Lévy-stable deviates with vanishing mean ($a = 0$) says that $|\Delta f(r;x_m)|^q = |\sum L(\alpha;0,c)|^q \stackrel{d}{=} r^{q/\alpha}|L(\alpha;0,c)|^q$. So, as long as $-1 < q < \alpha$, Lévy flight increments obey $\langle|\Delta f(r\ell;x_m)|^q\rangle \propto r^{q/\alpha}$; therefore $\zeta(q) = q/\alpha$. What if $q \geq \alpha$? In theory, the corresponding moments do not exist, so neither will the exponent $\zeta(q)$. In

[1] The symbol "$\stackrel{d}{=}$" means "identity in distribution."

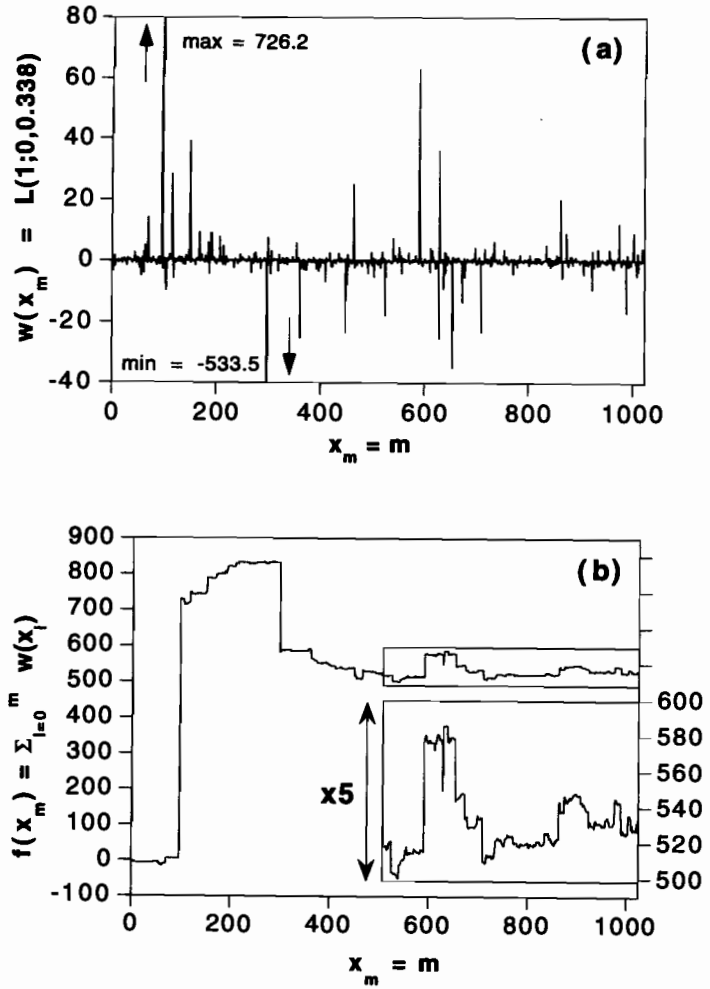


Figure A2: *Non-Ergodic Stationary and Nonstationary Processes.* (*a*) White Lévy-stable noise with the same order 1/2 moment as the Gaussian case in Fig. A1*a*; this data is stationary but non-ergodic (at least for moments of order $q \geq 1$). Symmetric Lévy-stable deviates with zero mean $L(\alpha;0,c)$ with a Lévy "index" $\alpha = 1$ and a scaling parameter $c = 0.338\cdots = \Gamma(3/4)^2/(\pi\sqrt{2})$ were used, obeying a Cauchy law: $\text{dProb}\{X \leq L(1;0,c) < X+dX\} = cdX/[\pi(c^2+X^2)]$. (*b*) Running sum of the data in panel (*a*), in other words, a "Cauchian" Lévy-flight; this data is nonstationary with increments that are stationary but not ergodic. The inset is a vertical zoom into the second half of the sample, showing more "jumps" of lesser magnitude.

practice, finite-sample estimates of $\langle|\Delta f(r;x_m)|^q\rangle$ become dominated by the most intense event in the w.n. $w(x_m) = L(\alpha;0,c)$. The corresponding value of x_m is straddled by exactly r of the $M-r$ ($> r$) increments included in the spatial averaging operation; this leads to $\langle|\Delta f(r;x_m)|^q\rangle \approx \max_m\{|w(x_m)|\}^q r/M \propto r$, hence $\zeta(q) = 1$ for $q > \alpha$. In summary, Lévy-stable processes are "operationally" multiscaling with $\zeta(q) = \min\{q/\alpha,1\}$; hence

$$H(q) = \min\{1/\alpha,1/q\},\ q > -1\ (0 < \alpha < 2). \qquad \text{(A6}b\text{)}$$

Here again we find $\beta = 2$ (operationally) although for radically different reasons than in Bm. To obtain Bm, we integrate ($\beta \rightarrow \beta+2$) a sample of w.n. ($\beta = 0$) with spectral density $E(k) \approx$ *constant* $< \infty$. In Lévy flights, spectral density (like variance) diverges: $E(k) \rightarrow \infty$ with sample size, via the prefactor. In a finite sample, $f(x)$ has a finite number of discontinuities, hence $\beta = 2$ as for Heaviside steps.[1]

A.2 Fractional Brownian Motions as Nonstationary Monofractal Functions

A.2.1 General Properties of Fractional Brownian Motion

Processes known as "fractional Brownian motion" (fBm) are, like Lévy flights, generalizations of Bm but in a different direction. They were introduced by Mandelbrot and van Ness (1968) to reproduce the correlation properties of otherwise Bm-like "steps" observed in numerous physical signals, for instance, turbulent velocity. Like standard Bm (uncorrelated steps), fBm is a nonstationary random function with stationary increments.

The noteworthy property of all scale-invariant nonstationary signals and models, fBm in particular, that describes incremental correlations is

$$\langle\Delta f(r;x+r)\Delta f(r;x)\rangle = [2^{2H(2)-1} - 1]\langle\Delta f(r;x)^2\rangle. \qquad \text{(A7)}$$

This follows directly from Eq. (16), expressed for r and $2r$ at $q = 2$, and the stationarity of the increments. So, at $H(2) = 1/2$ we retrieve Bm with characteristically uncorrelated increments in (A7). For $0 \le H(2) < 1/2$, we have *negatively* correlated increments: a jump up is more often than not followed by a jump down and vice-versa. This leads to a *less* nonstationary process than Bm, "anti-persistence" in Mandelbrot's (1983) words. For instance, Kolmogorov (1941) scaling in turbulent signals corresponds to $H(2) \approx 1/3$. For $1/2 < H(2) \le 1$, we have the opposite situation: *positively* correlated increments or "persistence" in Eq. (A7). If a jump tends to be followed by another in the same direction, then we have *more* nonstationarity than in Bm. Mandelbrot notes that the Earth's topography is reasonably well modeled by setting $H(2) \approx 0.75$.

The higher-order structure functions of fBm obey

$$H(q) = \zeta(q)/q \equiv H_2,\ 0 < H_2 < 1. \qquad \text{(A8)}$$

We recall from §4.2 in the main body that relation (A8) arises whenever the increments are distributed narrowly enough, Gaussian-style, that we can use the simplest dimensional

[1]The only difference in Fourier space between w.n. and a Dirac δ is the random phases; both have $\beta = 0$.

arguments to estimate higher-order moments from low-order ones: taking $q = 2$ as the standard, then $\langle|f(x+r)-f(x)|^q\rangle \approx \langle|f(x+r)-f(x)|^2\rangle^{q/2}$, $q > -1$. The definition of structure function exponents in Eq. (16) then leads to $\zeta(q) = \zeta(2)q/2 = qH(2) = qH_2$, hence (A8).

Equation (A8) shows that fBm belongs to the restricted class of *mono*fractal random functions; its only parameter relevant to scaling, H_2, is equal, in particular, to the mean Hölder exponent $H_1 = H(1)$. Figure A3*a* shows $\zeta(q)$ versus q for $H_1 = H_2 = 1/3$, $1/2$, $2/3$, hence an increasing degree of nonstationarity. The two extreme cases are also shown:

- $H_1 = H_2 \rightarrow 0$ is the "stationary" limit where increments are not only scale-invariant but scale-independent; and
- $H_1 = H_2 \rightarrow 1$ is the "differentiable" limit where increments become (almost) everywhere directly proportional to the scale.

In the remainder of this section, we describe the two major computational routes to fBm —one in physical space, the other in Fourier space— and a variant using running sums. The common denominator of these methods is to use only Gaussian random numbers and additions, hence the term "additive" is frequently used for this whole class of models. Sums of normal deviates are normally distributed, so all quantities involved are Gaussian —most importantly, $f(x)$ and $f(x+r)-f(x)$— therefore Eq. (A8) is verified by construction. Non-Gaussian multifractal models that violate (A8) were discussed above and will be again in §A.4, after surveying the prerequisite singular measures in s §A.3.

A.2.2 Synthesis in Physical Space (Recursive Mid-Point Displacement)

The simplest algorithm for making fBm is known as "mid-point displacement," cf. Peitgen *et al.* (1988). For procedures in physical space, it is generally convenient to use units of length where the grid constant is unity and we take a power of 2 plus one grid size:

$$M = L_n + 1 = 2^n + 1,\ n > 0. \qquad \text{(A9)}$$

To determine $f_n(x_j)$ for $j = 0,\ldots,2^n$, we first set[1]

$$f_n(0) = f_0 \qquad \text{(A10)}$$

$$f_n(L_n) = \begin{cases} N(f_0,\sigma_0) \text{ where } \sigma_0 = 2^{(1-H_2)}-1, \text{ or} \\ f_n(0) \text{ for cyclical boundary conditions} \end{cases} \qquad \text{(A11)}$$

This completes the maximum scale $r_0 = L_n$. We then proceed recursively to smaller and smaller scales,

$$r_i = r_{i-1}/2 = r_0/2^i = 2^{n-i}\ (i = 1,\ldots,n), \qquad \text{(A12)}$$

using a decreasing sequence of standard deviations,

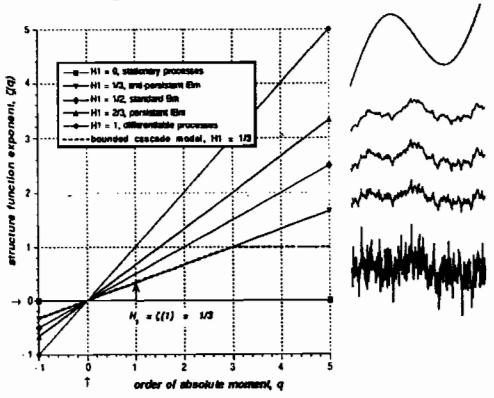
$$\sigma_i = \sigma_{i-1}/2^{H_2} = \sigma_0/2^{iH_2}\ \ (i = 1,\ldots,n). \qquad \text{(A13)}$$

New points are generated by averaging existing ones and adding random displacements:

$$f_n(x_j+r_i) = [f_n(x_j+2r_i)+f_n(x_j)]/2 + N(0,\sigma_i),\ x_j = 2jr_i\ (j = 0,\ldots,2^i-1). \qquad \text{(A14)}$$

[1]The value of $f_n(0)$ in (A10) is arbitrary but deterministic, hence the patently nonstationary nature of fBm.

Results for five cases are presented in Fig. A3*b*, illustrating more and more nonstationarity: $H_2 = 0$ ("1/*f*" noise, cf. next section); $H_2 = 1/3$ (anti-persistence, as in turbulent signals); $H_2 = 1/2$ (Bm); $H_2 = 2/3$ (persistence, as in topography); and $H_2 = 1$ (a noiseless linear trend). For the generalization to 2D, we refer to Peitgen *et al.* (1988).



(a) (b)

Figure A3: *Fractional Brownian Motions in 1D, with More or Less Nonstationarity.* (*a*) *q*th-order structure function exponents $\zeta(q) = qH_1$. (*b*) Three different fBm's ($H_1 = H_2 = 1/3$, 1/2, 2/3, by increasing degree of nonstationarity) and two limiting cases: $H_1 = 0$ ("1/*f*noise," border-line stationarity) and $H_1 = 1$ (differentiability, maximum nonstationarity on this scale).

A.2.3 Synthesis in Fourier Space (Power-Law Filtering)

Another way of generating fBm is with low-pass power-law filtering in Fourier space, a procedure also known as "fractional integration" that we will invoke again further on to generate multifractal signals.[1]

Following Voss (1983), we start with Gaussian white noise on a grid of size 2^n:

$$w_n(x_j) = N(0,1), j = 0,\ldots,2^n-1. \tag{A15}$$

This trivial stochastic process made of normal deviates, completely uncorrelated from one grid point to the next, is scale-invariant by construction and characterized by a spectral exponent

$$\beta_w = 0. \tag{A16}$$

We now want to incorporate correlations, leading to

$$\beta = 2H_2 + 1. \tag{A17}$$

This is easily done by power-law filtering in Fourier space, where it is convenient to set

$$L = 2^n \ell_n = 1. \tag{A18}$$

[1]Pearson (1990) surveys fractional integration, as defined in the mathematical literature, which is somewhat different from ours.

After obtaining[1]

$$\begin{cases} \tilde{f}_n(0) = \tilde{w}_n(0) \approx 0 \\ \tilde{f}_n(k) = \tilde{w}_n(k) \times |k|^{-(H_2+1/2)}, k = \pm 1,\ldots,\pm 2^{n-1}, \end{cases} \tag{A19}$$

we compute $f_n(x_j)$ by inverse Fourier transform for $x_j = j\ell_n$ ($j = 0,\ldots,2^n-1$). The exponent $H_2+1/2$ ($0 < H_2 < 1$) in Eq. (A19) is naturally the "order of the fractional integration." Generalization to 2D is straightforward.

Equation (A17) still applies if we take $H_2 < 0$ hence $\beta < 1$, or $H_2 > 1$ hence $\beta > 3$, in (A19). The latter choice yields nonstationary random functions with *non*stationary increments, varying so weakly in the small scales that they are everywhere continuous and almost everywhere differentiable. So, increments become proportional to distance between points of interest: $\langle |f(x+r)-f(x)| \rangle \approx \langle [f(x+r)-f(x)]^q \rangle^{1/q} \propto r$, leading to $H(q) \equiv 1$. We now discuss the consequences of the former choice.

A.2.4 Synthesis from Stationary Gaussian Scaling Noises

What if we take $H_2 < 0$ in the recipe contained in Eq. (A16)? The recursive mid-point displacement procedure in §A.2.3 can also be applied for negative[2] H_2. However, the outcome for $H_2 < 0$ and $H_2 > 0$ have radically different properties.

- If $H_2 < 0$, the corresponding scale-invariant noise has an energy spectrum $E(k) \propto k^{-\beta}$ with $\beta < 1$; there is finite energy (i.e., variance) at large scales ($k \to 0$) but divergence at small ones ($k \to \infty$) —a so-called "ultra-violet" catastrophe occurs. This means that small-scale singularities will develop whereas large-scale properties will remain relatively well-behaved.
- For $H_2 > 0$ ($\beta > 1$), the opposite occurs —an "infra-red" catastrophe. We can therefore expect large fluctuations at scales comparable to the computational domain, but relatively small ones at the pixel scale.

Only for $H_2 < 0$ is there a well-defined, monotonically decreasing power-law autocorrelation function:[3] $\langle f(x+r)f(x) \rangle \propto r^{-(1-\beta)} = r^{-2|H_2|}$, hence stationarity in the broad-sense. For $0 < H_2 < 1$, a power-law structure function for $q = 2$ follows from a power-law energy spectrum $E(k)$: $\langle [f(x+r)-f(x)]^2 \rangle \propto r^{2H_2}$, hence (broad-sense) stationary increments.

We will refer to these models with $H_2 < 0$ as "Stationary Gaussian Scaling Noises" (or SGSNs). One-dimensional fBm with spectral exponent $1 < \beta < 3$ can be obtained by standard integration (i.e., running sums in physical space) of SGSNs with $\beta' = \beta-2$ such that $|\beta'| < 1$. These SGSNs are first obtained by fractional integration of Gaussian w.n. at order $\beta'/2 = H_2+1/2$, ranging from $-1/2$ to $1/2$ ($-1 < H_2 < 0$). We note that, for the anti-persistent case ($1 < \beta < 2$), the corresponding SGSNs results in effect from a fractional

[1]Note that we leave the phases $\tan^{-1}(\mathrm{Im}[\tilde{w}_n(k)]/\mathrm{Re}[\tilde{w}_n(k)])$ of the Fourier components unchanged. However, to retrieve Bm exactly (as defined in §A.1.1), we need to change the phases by exactly $\pi/2$: $\tilde{f}_n(k) = \tilde{w}_n(k)/(ik)$ will yield a realization with $f_0 = \tilde{w}_n(0)L \approx 0$.

[2]The choice $H_2 = -1/2$ leads to a rather convoluted way of generating uncorrelated values at every pixel. In contrast, this choice implies a null operation in the Fourier-based recipe in Eq. (A19).

[3]This requires $0 < \beta < 1$ ($-1/2 < H_2 < 0$), otherwise ($H_2 < -1/2$) there are anti-correlations from one grid-point to the next; thus $\langle f(x+r)f(x) \rangle < 0$, and a power-law parameterization is invalid.

"differentiation" of Gaussian white noise, the order of "integration" $\beta'/2$ being negative, between 0 (for Bm) and $-1/2$ (for "$1/f$" noise).

A.3 Cascades Leading to Singular, Stationary, Mono- and Multifractal Measures

Calibration of singularity analysis algorithms calls for a different class of models than described above. The main requirement is that the model can be read as a "measure," a non-negative[1] field that we will denote generically by $\varepsilon(x)$. This excludes the models discussed up to now. In practical data analysis applications, the nonstationary data-stream can be used to generate a measure, typically by taking small- but finite-scale gradients then their absolute values or squares. We can do this for the models presented in section A.2 but the results are quite disappointing, the small-scale gradients fields of Gaussian models being at best weakly variable (narrowly distributed).

Let $f(x)$ be a nonstationary *additive* process based on $w(x)$, a Gaussian-type noise such as an SGSN. This gives $\varepsilon(x) = |w(x)|$ which varies less than $w(x)$ itself. Being essentially a running mean, the coarse-grained measure $\varepsilon(r;x) = (1/r)\sum_{x'=x}^{x+r-1}|w(x')|$ will be statistically independent not only of x, stationarity oblige, but also of r over a large range of values because of the strong ergodic property of Gaussian-like $w(x)$. This translates to $K(q) \equiv 0$ in Eq. (20), apart from finite sampling effects (Aurell *et al.* 1992). The same is true for models of $w(x)$ with power-law tails, except that moments exist only up to some critical order q_D: $K(q) \equiv 0$ for $q < q_D$; for instance, Lévy-stable variables yield $q_D = \alpha < 2$.

To obtain models with $K(q) \not\equiv 0$, we must leave the realm of *additive* models and generate measures with *multiplicative* cascades. In the same way as fBm's are monofractal random functions $f(x)$ with multifractal counterparts (some of which are described in Section A.4), there are mono- and multifractal measures. The literature on multifractal measures is vast and increasing at a rapid pace. In this survey, we will start with some well-known monofractal cases in 1D and 2D (§A.3.1); straightforward generalizations of these (§A.3.2) lead to multiscaling in 1D (§A.3.3–4) and 2D (§A.3.5).

Mono- and multifractal cascade models come in two distinct flavors: the "canonical" (§A.3.3) and the "microcanonical" (§A.3.4–5). The latter, having the same singularity properties for every realization (ergodicity in the sense of singular measures), are particularly useful for the purposes of calibration; the former are arguably more realistic models for geophysical fields (Schertzer and Lovejoy 1987, Gupta and Waymire 1993). However, they are always non-ergodic at some level of confidence and can sometimes exhibit the interesting statistical phenomenon of divergence of higher-order moments. If ignored, this last characteristic can affect a calibration procedure.

A.3.1 Cantor's, Dirac's and Other Monofractal Measures

Consider the example of Cantor's measure, supported entirely by Cantor's famous set (Fig. A4*a*) which has fractal dimension $D_f = \log 2/\log 3 = \log_3 2 = 0.631\cdots < d = 1$.

[1]Mathematically, "measures" can be thought of as "generalized" functions in the sense of Dirac (i.e., defined only under integrals) that are furthermore non-negative. The sum of integrals over disjoint sets is therefore equal to the integral over their union.

Cantor's measure is generated by starting with a uniform distribution, $\varepsilon_0(x) = 1$ on [0,1] which is then divided into 3 equal parts. The middle[1] third is emptied of its measure —this cell (or "eddy") is now "dead"— and its mass is uniformly redistributed between its neighbors where $\varepsilon_0(x)$ therefore becomes $\varepsilon_1(x) = \varepsilon_0(x)\times(1+1/2) = 3/2$. We have thus required the spatial average to remain constant. After n steps, the scale is $r_n = 3^{-n}$, there are 3^n cells in all, only 2^n of which carry a measure $\varepsilon_n(x) = (3/2)^n$, elsewhere $\varepsilon_n(x) = 0$. Although deterministic, one can still define the spatial statistics of this measure: $\langle\varepsilon(r_n;x)^q\rangle = (2/3)^n\times(3/2)^{nq} = [(1/3)^n]^{(1-q)\log_{1/3}(2/3)}$. Equation (20) in the main text then yields

$$K(q) = [1-\log_3 2](q-1),\ q > 0. \tag{A20a}$$

For $q = 0$ we have $K(q) = 0$ (by definition) and, for $q < 0$, we find $\langle\varepsilon(r_n;x)^q\rangle = \infty$ (formally, we can set $K(q) = \infty$ here) because the empty events, $\varepsilon(r_n;x) = 0$, dominate. We note that the factor $1-\log_3 2$ in (20*a*) is the codimension $d-D_f$ of Cantor's set.

An example embedded in 2D is easily generated as follows. Start with a uniform distribution, $\varepsilon_0(x,y) = 1$, on the unit square; then divided it into 4 equal parts; one sub-square is picked at random[2], and its measure (1/4 of total) is set to 0; the measure in the neighboring cells is boosted proportionately, becoming $\varepsilon_1(x,y) = \varepsilon_0(x,y)\times(1+1/3) = 4/3$. After n steps, the scale of interest is $r_n = 2^{-n}$; there are 4^n cells in all with only 3^n of them containing a non-vanishing measure, $\varepsilon_n(x,y) = (4/3)^n$. Spatial averaging leads therefore to $\langle\varepsilon(r_n;x)^q\rangle = (3/4)^n\times(4/3)^{nq} = [(1/2)^n]^{(1-q)\log_{1/2}(3/4)}$, hence

$$K(q) = [2-\log_2 3](q-1),\ q > 0. \tag{A20b}$$

The support of this measure is a fractal set of codimension $-K(0^+) = 2-\log_2 3$, hence a fractal dimension $D_f = 2+K(0^+) = \log_2 3 = 1.585\cdots < d = 2$. as can by verified by direct box-counting methods.[3]

The simplest random monofractal measure is probably a Dirac δ-function positioned at a random point x^* on the unit segment in $d = 1$: $\varepsilon(x) = \delta(x-x^*)$, $0 < x^* < 1$; we can also consider the unit square in the $d = 2$ case: $\varepsilon(x,y) = \delta(x-x^*)\delta(y-y^*)$, $(x^*,y^*) \in [0,1]^2$. These measures are entirely concentrated onto a single point; in other words the fractal dimension of their support is $D_f = 0$, hence a codimension $d-D_f = d$. Let us estimate the statistical moments of the coarse-grained measure $\varepsilon(r;x)$ contained in the sub-interval $[x,x+r)$ or sub-domain $[x,x+r)\otimes[y,y+r)$, $0 \le r \le 1$, according to Eq. (14): taking x (and, if necessary, y) at random, we find $\varepsilon(r;x) = 1/r^d$ with probability r^d and 0 otherwise. Therefore, we have $\langle\varepsilon(r;x)^q\rangle = r^d/r^{dq} = r^{(1-q)d}$; hence, from Eq. (20):

$$K(q) = (q-1)d,\ q > 0. \tag{A20c}$$

All of the above formulas (A20*a,b,c*) can be parameterized as

$$K(q) = (q-1)C_0,\ q > 0, \tag{A21}$$

[1]The "middle-third" convention leads to a deterministic measure. The spatial statistics are unchanged by picking the next dead cell at random.

[2]If the same (e.g., upper-left) sub-square is picked each time, the limiting set is akin to Sierpinski's triangle.

[3]Total number of boxes of size $r = 1/2^n$ ($n \ge 0$) in unit square: $N_t = (2^n)^2 = (1/r)^d$ ($d = 2$). Number needed to cover fractal set: $N_s = 3^n \propto (1/r)^{D_f}$. Fractal dimension of set: $D_f = \log N_s/\log(1/r) = \log_2 3$ (QED).

where $C_0 = d - D_f > 0$ is the codimension of the support. The above measures are "monofractal" in the sense that their supports are nontrivial fractals; however, on these sparse subsets of space, they are uniformly distributed. Why their exponents $K(q)$ are always in the linear relationship spelled out in Eq. (21) can be traced to this uniformity.

Another common feature of the above monofractal models is that they were constrained, largely for tutorial purposes, to have the same statistics for every realization (and indeed at every cascade step); in particular, the total measure is conserved. Such models are called "microcanonical" (Mandelbrot, 1974). "Canonical" counterparts of these models —where only the probability of killing a cell is prescribed, irrespective of what happens to its neighbors— have been proposed to model 1D transects of the kinetic energy dissipation field in fully-developed 3D turbulence (Novikov and Stewart, 1964; Mandelbrot, 1974; Frisch *et al.*, 1978). Indeed there is no reason to require a 1D sample of a 3D field to obey a conservation law that applies at best to the whole volume of the system. The observed statistics of turbulent flows are best reproduced by models with C_1 values in the range 0.2–0.3. Following the nomenclature of Frisch *et al.* (1978), we refer to these monoscaling canonical cascade models collectively as "beta" models.

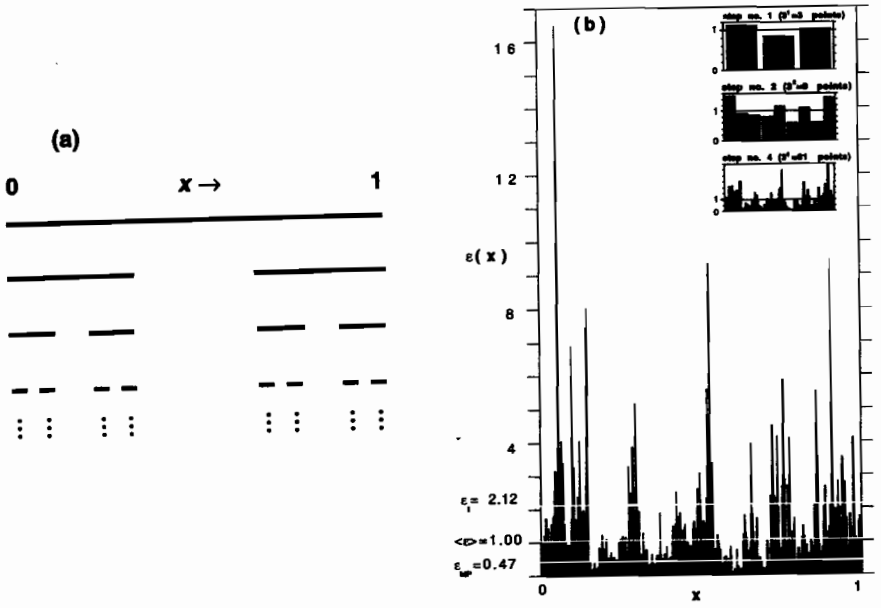


Figure A4: *Two Multiplicative Cascades in 1D with branching ratio* $\lambda = 3$. (*a*) First 3 cascade steps in the construction of Cantor's deterministic measure, with $D(q) \equiv \log_3 2 = 0.63\cdots$. (*b*) A random log-normal measure, cascade steps # 1, 2, 4 (inset) and 6, with $D(q) = d - C_1 q$ where $d = 1$, $C_1 = 0.11\cdots$ and $q < d/C_1$.

A.3.2 Canonical Random Measures and Divergence of Higher-Order Moments

There is no reason to limit ourselves to uniform measures. A simple way of obtaining non-uniform measures is to allow the random multiplicative "weights" used implicitly in the above cascade algorithms to differ from 0 or a constant. In general, we can envision a "turbulent" cascade process in $d = 1$ or 2 dimensions that proceeds by divisions into λ^d "sub-eddies" of equal size at each step; the integer λ is known as the "branching ratio." After n steps the scale is $r_n = \lambda^{-n}$, we have

$$\varepsilon_n = \prod_1^n W_i, \tag{A22a}$$

requiring only that

$$\langle W \rangle = 1. \tag{A22b}$$

Redefining temporarily $K(q)$ as $-\ln\langle \varepsilon_n{}^q \rangle / \ln r_n$ —as opposed to the parameterization in the main text's Eq. (20) of the integrals (measures) defined in Eqs. (14–15)— it is easy to see that

$$K(q) = \frac{\ln\langle W^q \rangle}{\ln\lambda} = \log_\lambda \langle W^q \rangle, \ \{q \in \Re; \langle W^q \rangle < \infty\}. \tag{A23a}$$

Defined in this way, $K(q)$ inherits all of the analytical properties of cumulant-generating function[1] for the random variables $\ln W$. In particular, $K(q)$ is convex with $K(0) = 0$. We also know that $K(1) = 0$ due to (A22*b*).

Returning to the original definition of $K(q)$ in relation to the measures in Eq. (20) is quite involved mathematically (Mandelbrot, 1974; Kahane and Peryière, 1976; Schertzer and Lovejoy, 1987; Gupta and Waymire, 1993). However, the only difference is ultimately that the trivial condition on q in Eq. (A23*a*) is replaced by

$$\{q < q_D; \frac{K(q_D)}{q_D - 1} = d\}. \tag{A23b}$$

For $q \geq q_D$, the moments $\langle \varepsilon(r,x)^q \rangle$ are divergent. In practical data analysis applications, the symptom of a diverging moment is that its estimate is unstable, being dominated by the single largest event. The intensity of this event will depend critically on the sample size (in this case, grid size and number of realizations); we refer to Schertzer and Lovejoy (1992) for further details.

In summary, we can view the monofractal models of §A.3.1 as a limiting case in a continuum of models. The codimension of the measure's support is[2]

$$C_0 = -K(0^+). \tag{A24}$$

If, as in monofractal cases, $C_0 > 0$ then it is arguably the most important parameter of the model since it defines geometrically the concentration of the measure on a sparse subset

---

[1] The characteristic function of a random variable ξ is defined as $\phi(t) = \langle \exp[it\xi] \rangle$, i.e., the Fourier transform of its pdf; the cumulant-generating function (or "2nd characteristic function") is $\ln\phi(t)$. Taking $\xi = \ln W$ and $it = q$, we have $K(q) = \ln\phi(q/i)/\ln\lambda$. For imaginary arguments, $\phi(t)$ is real and convex (Feller 1971).

[2] For $q = 0^+$ everywhere $\varepsilon(r,x) > 0$ we have $\varepsilon(r,x)^q = 1$, otherwise (i.e., $\varepsilon(r,x) = 0$), we have $\varepsilon(r,x)^q = 0$.

of space. In many cases however (cf. next sub-section), the probability of drawing exactly null weights is vanishingly small; this leads to $K(0^+) = K(0) = 0$, hence $C_0 = 0$.

In situations where $C_0 = 0$ (i.e., the measure is supported by all of space), the next simplest way of quantifying its degree of concentration (hence intermittency) is the information codimension:

$$C_1 = K'(1). \tag{A25}$$

This quantity can be obtained directly from the weights since Eqs. (A23*a*) and (A25) yield $K'(1) = \langle W \log_\lambda W \rangle$. For monofractal models, the linear formula for $K(q)$ in Eq. (A21) yields $K'(1) = K(0^+)$, hence $C_0 = C_1$. We also note that, $W\log_\lambda W$ being a convex function of W, Jensen's (1906) inequality tells us that $C_1 = \langle W \log_\lambda W \rangle \geq \langle W \rangle \log_\lambda \langle W \rangle = 0$.

Yet another way of parameterizing intermittency uses the correlation codimension

$$C_2 = K(2). \tag{A26}$$

This 2nd-order statistic is the preferred choice in the turbulence literature, where C_2 is in fact referred to as the "intermittency parameter" (and denoted "μ"). It determines the measure's spectral exponent[1]

$$\beta_\varepsilon = 1 - K(2) < 1. \tag{A27}$$

The inequality follows from $K(q)$'s convexity and $K(1) = 0$ which imply $K(2) > 0$. This establishes the stationarity of cascade processes according to the criterion in Eq. (28*a*).

A.3.3 The Canonical Log-Normal Model, Stationarity in Presence of Intermittency

Consider a specific example: we use the Cantor measure construction ($\lambda = 3$) with log-normal rather than Bernoulli weights. Instead of $W_i = 0$ (Prob 1/3) or 3/2 (Prob 2/3), we take:

$$W_i = \exp[N(\mu,\sigma)] = \exp[\mu + \sigma N(0,1)] \quad (i = 1,\ldots,n), \tag{A28a}$$

with

$$\mu = -\frac{\sigma^2}{2}, \tag{A28b}$$

to properly normalize the cascade, $\langle W \rangle = 1$.

In Fig. A4*b*, we illustrate the first steps of the construction using $\sigma = 0.4$. Notice the development of multiple singularities. At the same time, most of $\varepsilon_n(x)$'s values become smaller as n is incremented. Indeed, at constant x, $\ln\varepsilon_n(x) = \sum_0^n \ln W_i = \sum_0^n [\mu + \sigma N(0,1)]$ is executing a random walk as a function of (discrete) "time" n, like the Bm described in §A.1.1 but with a systematic drift in the negative direction, due to (A28*b*).

Log-normal cascades were first introduced in turbulence theory by Kolmogorov (1962) and Obukhov (1962) to account for the effects of intermittency in the kinetic energy dissipation rate at high Reynolds numbers; σ around 0.7 was found to fit the data. We recall that, like SGSNs (§A.2.4), the model in Fig. A4*b* exemplifies stationarity according to criterion (28*a*) in the main text, cf. Eq. (A27). However, in sharp contrast with SGSNs, these cascades are highly non-Gaussian, not unlike the Lévy-stable w.n.

[1]This follows from $\langle \varepsilon(x)\varepsilon(x+r) \rangle \sim r^{-K(2)}$ (Monin and Yaglom 1975, pp. 618–620) which is in Fourier duality with the energy spectrum (Wiener–Khinchin theorem) $E_\varepsilon(k) \sim k^{-\beta_\varepsilon}$ with $\beta_\varepsilon + K(2) = 1$.

described previously. This is traceable to the multiplicative nature of the construction. In contrast with the white Lévy-stable noise, cascades have more interesting correlation properties due to the recursive nature of their construction.

Equations (A23*a*) and (A28*a,b*) yield[1]

$$K(q) = \frac{\sigma^2}{2\ln\lambda} q(q-1), \tag{A29}$$

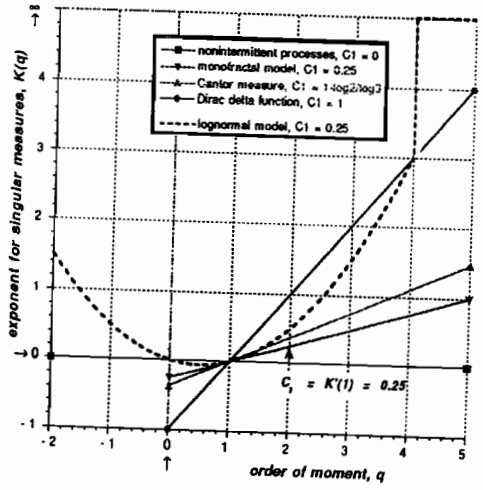
Using (A25), this can be rewritten[2]

$$K(q) = C_1 q(q-1),\ q < q_D = d/C_1, \tag{A30}$$

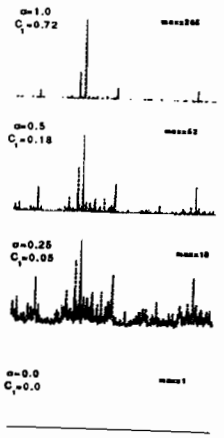
by identifying C_1 with $\sigma^2/(2\ln\lambda)$. The condition on q in Eq. (A30) accounts for the divergence of moments, as predicted by Eq. (A23*b*).

In Fig. A5*a*, we have plotted $K(q)$ for a log-normal model in $d = 1$ with $C_1 = 0.25$; notice the divergence of moments, formally $K(q) = \infty$, for $q > q_D = 4$. For comparison, we have also plotted $K(q)$ for two of the monofractal models discussed previously, Dirac δ's and Cantor measures. Figure A5*b* shows a sequence of log-normal cascade models with increasing degrees of intermittency: $\sigma = 0$ ($C_1 = 0$); $\sigma = 0.25$ ($C_1 \approx 0.05$); $\sigma = 0.5$ ($C_1 \approx 0.18$); and $\sigma = 1.0$ ($C_1 \approx 0.72$). Notice the increasing concentration and spikiness.

Figure A5: *Log-Normal Cascade Models in 1D, with More-or-Less Intermittency.* (*a*) Exponents $K(q)$ for a log-normal multifractal case and two monofractal cases; the onset of divergence for $q > 4$ is indicated for the 1D log-normal model with $C_1 = 0.25$. (*b*) Four cascades with $\langle \varepsilon \rangle = 1$ that illustrate an increasing degree of intermittency: $C_1 = 0.0$, 0.05, 0.18, and 0.72. Models with $C_1 > 1$ would be "degenerate" in 1D, meaning almost empty in most realizations but a huge peak would occur now and then.



(a)



(b)

[1]The characteristic function for normal deviates is $\langle \exp[itN(\mu,\sigma)] \rangle = \exp[i\mu t - t^2(\sigma^2/2)]$, where we set $t = q/i$.

[2]Schertzer and Lovejoy (1987) propose a class of log-Lévy cascade models parameterized by $\alpha \in [0,2]$. They run the gamut between beta-models in Eq. (A21) and log-normal models in Eq. (A29): if $\alpha \neq 1$, $K(q) = C_1(q^\alpha - q)/(\alpha - 1)$, and $K(q) = C_1 q \ln q$ for $\alpha = 1$.

A.3.4 The Microcanonical Log-Binomial "p-model"

Meneveau and Sreenivasan's (1987*b*) one-dimensional "*p*-model" is a microcanonical alternative to the canonical log-normal model. Here $\lambda = 2$ and the weights are

$$W_i = 1 \pm (1-2p) = \left\{ \begin{array}{ll} W_- = 2p & (\text{Prob} = 1/2) \\ W_+ = 2-2p & (\text{Prob} = 1/2) \end{array} \right\}, \; 0 < p < 1/2 \; (i = 0,\ldots,n\text{-}1) \qquad (A31a)$$

in one sub-eddy, and

$$W'_i = 2 - W_i = \left\{ \begin{array}{l} W_+ \\ W_- \end{array} \right\} \qquad (A31b)$$

in the other, to conserve the total measure (cf. Fig. A6). The microcanonical nature of this model means that every sequence of ±'s yields the same binomially distributed values for $\varepsilon_n(x)$, $W_-^m W_+^{n-m}$ ($m = 0,\ldots,n$) with probability $\binom{n}{m}/2^n$; only their order of occurrence changes. Such predictability —exact ergodicity in the sense of singular measures— in a calibration procedure for singularity analysis is obviously desirable.

The choice of weights in Eq. (A31*a*) yields

$$K(q) = \log_2[W_-^q/2 + W_+^q/2] = \log_2[(2p)^q + (2\text{-}2p)^q] - 1, \qquad (A32)$$

for[1] $q \in \Re$. Exponents of special interest are

- the information dimension: $D_1 = 1-C_1 = -[p\log_2 p + (1-p)\log_2(1-p)]$ from $q = 1$; and
- the spectral exponent: $\beta_\varepsilon = 1-\log_2[1+(1-2p)^2]$ from $q = 2$.

The latter is plotted as a function of p in Fig. A7*a* where the inset illustrates a typical realization. The sample in question was generated using $p = 0.35$, as suggested by Meneveau and Sreenivasan's (1987*a*) measurements of the dissipation field in turbulent flows; its more important characteristics are $D_1 \approx 0.65$ ($C_1 \approx 0.35$), and $\beta \approx 0.88$. The two limiting cases are familiar: $p \to 0^+$ yields randomly placed δ-functions, $p \to 1/2^-$ (or $\sigma \to 0^+$ for the log-normal case) weak *1/f*-type fluctuations from the unitary mean value.

A.3.5 The "p(3)-model", A 2D Generalization of the "p-model"

The most general microcanonical cascade with $\lambda = 2$ in $d = 2$ calls for three parameters that we will denote

$$0 \le p_1, p_2, p_3 \le 1/2. \qquad (A33)$$

They will be used to shift mass in the same way as in the *p*-model but in both horizontal (E–W) and vertical (N–S) directions as well as between the two diagonals:

$$\left\{ \begin{array}{ll} \text{E} \leftrightarrow \text{W:} & 1 \pm (1-2p_x) \\ \text{N} \leftrightarrow \text{S:} & 1 \pm (1-2p_y) \\ \text{NE/SW} \leftrightarrow \text{NW/SE:} & 1 \pm (1-2p_d) \end{array} \right\}, \qquad (A34)$$

[1] We have $K(q) \approx q\log_2 W_\pm - 1$ for $q \to \pm\infty$. From Eq. (A23*b*), and knowing that $W_+ = 2-2p < 2$, the large q limit tells us that all moments converge for microcanonical models. Not only they converge but (by construction) their estimates for every realization are identical, even for a single cascade step.

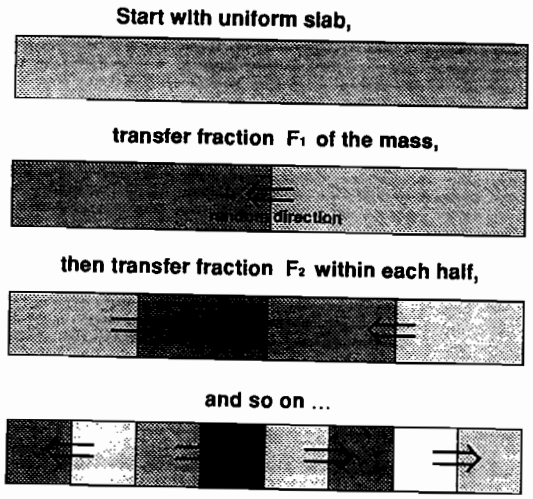


Figure A6: *Genesis of a Microcanonical Cascade in 1D.* Three steps are illustrated. The total mass is conserved at each step, implying anti-correlated multiplicative weights in each sub-cell: $W_i + W_i' = 2$, hence $W_i = 1 \pm F_i$ and $W_i' = 1 \mp F_i$.

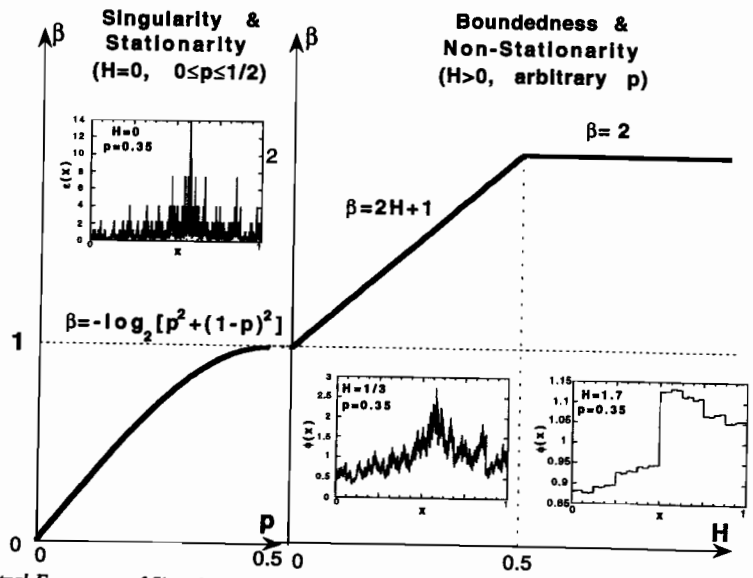


Figure A7: *Spectral Exponents of Singular and Bounded Cascades.* (*a*) Stationary but intermittent case $H = 0$, reverting to Meneveau and Sreenivasan's (1987) "*p*-model" for the kinetic energy dissipation rate that occurs at the Kolmogorov scale in fully developed turbulent flows; this model is unbounded ("singular") in the limit of many cascade steps. (*b*) Nonstationary $H > 0$ generalization proposed by Cahalan *et al.* (1990, 1994) to model cloud optical depth variability in stratocumulus decks that are observed to have power-law spectra with $\beta \approx 5/3$.

where $\{p_x, p_y, p_d\}$ is one of the 3! = 6 permutations of $\{p_1, p_2, p_3\}$, with equal probability and random ±. After the three transfers, the individual combined weights are

$$W = \prod_{\theta \in \{x,y,d\}} W_\theta = \prod_{k=1}^{3} [1 + s_k(1-2p_k)], \tag{A35}$$

where $\{s_k; k=1,2,3\} = \{s_1, s_2, s_3\} = \{\pm, \pm, \pm\}$, one of the 8 equally probable combinations. These weights fall between

$$W_\pm = \min\{W\} = \prod_{k=1}^{3} [1 \pm (1-2p_k)] \tag{A36}$$

For this choice of W's,

$$\langle W^q \rangle = \frac{1}{8} \sum_{\{\pm,\pm,\pm\}}^{8\ \text{combos}} \prod_{k=1}^{3} [1 \pm (1-2p_k)]^q, \tag{A37}$$

for $q \in \Re$, hence $K(q)$ from Eq. (A23*a*). If $p_1 = p < 1/2$ and $p_2 = p_3 = 1/2$, we have only two equally probable weights $W_\pm$, and Eq. (A32) is retrieved.

A.4 Going from Stationary Measures to Nonstationary Multifractal Functions

It is important to have the option of multiscaling as well as monoscaling (section A.2) in the realm of nonstationary functions with stationary increments. We therefore need algorithms that generate fields where structure-function analysis yields a nonlinear $\zeta(q)$, equivalently, a non-constant $H(q)$. There are many well-documented methods for generating stationary multifractal measures using multiplicative cascades; in contrast, the literature on specific ways of generating nonstationary functions with multiscaling structure functions is relatively small. We present two procedures here (Schertzer and Lovejoy, 1987; Cahalan *et al.*, 1990) and refer to Viscek and Barabasí (1991), Arnéodo *et al.* (1992), Benzi *et al.* (1993), and Sykes *et al.* (1996) for the few others we are aware of.

A.4.1 "Bounded" Cascades

One route from multiplicative cascades to nonstationary multiscaling processes was charted by Cahalan *et al.* (1990) for the purposes of modeling the internal structure of marine stratocumulus. Their model builds on Meneveau and Sreenivasan's (1987) p-model, calling for a power-law decay of the variance of the weights as the multiplicative cascade proceeds to smaller and smaller scales, $r_i = L/2^i$ $(i = 0,\ldots,n)$. Explicitly, we take

$$W_i = 1 \pm (1-2p)(r_{i-1}/L)^{-H} \text{ (equal Prob ±)}, 0 < p \le 1/2, 0 \le H \le \infty, \tag{A38a}$$

$$W'_i = 2 - W_i, \tag{A38b}$$

for $i \ge 1$. Figure A7*b* shows the effect of H on the spectral exponent: $\beta = \min\{2H,1\}+1$ (Cahalan *et al.*, 1994); the inset shows examples for $H = 1/3, 1.7$.

Marshak *et al.* (1994) show that the resulting model, $f_n(x)$, is multifractal in the sense that higher-order structure functions: $\langle |f_n(x+r)-f_n(x)|^q \rangle$ goes as $r^{qH(q)}$ with

$$H(q) = \min\{H, 1/q\}. \tag{A39}$$

Interestingly, p does not appear in Eq. (A39) which determines the scaling of 2-point statistics of all orders; Cahalan *et al.* (1994) investigate its role in 1-point statistics, hence prefactors in Eq. (16) of the main text. Clearly, H controls the degree of nonstationarity. The p-model is retrieved in the "stationary" limit $H = 0$. At $H \to \infty$, we find the "most nonstationary" in this continuum of models, Heaviside steps of height $2(1-2p)$ at $x = L/2$.

Figure A8 illustrates two important properties of bounded cascades which are shared by other multiscaling random functions (as well as fBm's, their monoscaling counterparts): stochastic continuity and self-affinity. Two identical sequences of successive horizontal zooms are illustrated; the difference is only that on the l.h.s. the vertical scale is held constant and on the r.h.s it is rescaled by a given factor at each zoom. On the l.h.s. we see the amplitude of the increments decrease dramatically with scale (continuity property). However, on the r.h.s. we see that zooms onto different portions of the graph of $f_n(x)$ are statistically similar: the graph is self-affine with fractal dimension $D_{graph} = 2-H_1 = 5/3$ in this $H_1 = H = 1/3$ case.

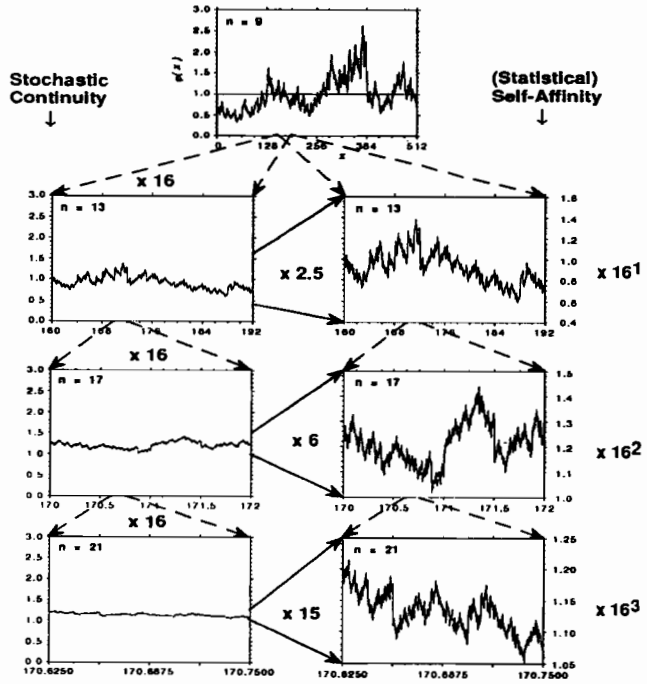


Figure A8: *Stochastic Continuity and Self-Affinity of a 1D Bounded Cascade Model.* The sequence of zooms on the **l.h.s.** (without vertical rescaling) shows how smaller scales lead to smaller increments; this is called "stochastic continuity." The **r.h.s.** sequence (with vertical rescaling) yields three graphs that are statistically indistinguishable from the original at the top; this property is "self-affinity."

Figure A9 shows an example of a 2D generalization of the above bounded cascade model starting with the "$p(3)$-model." The factors $1-2p_k$ (k = 1,2,3) in Eqs. (A34–36) are multiplied by the variance moderating factor $(r_{i-1}/L)^{-H}$, $i = 1,\ldots,n$, as in Eq. (38a). This type of model was used by Cahalan (1994), Marshak *et al.* (1995a) and Davis *et al.* (1996b) in radiative transfer computations.

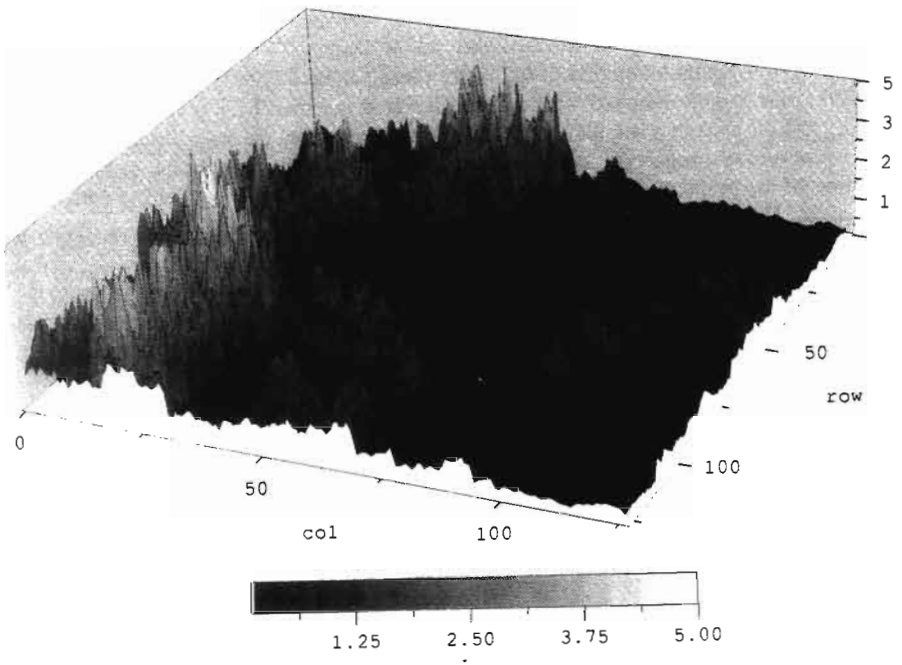


Figure A9: *Perspective View of a 2D Bounded Cascade Model.* The numerical values of parameters are H = 1/3 and $p_k = 0.42+(k-1)\times 0.03$ in a 7-step process.

A.4.2 Random Devil's Staircases

The simplest way of obtaining a function from a multifractal measure is to take its indefinite integral:

$$f_\varepsilon(x) = \int_0^x \varepsilon(x')dx',\ 0 \le x \le L. \tag{A40}$$

Since $\varepsilon(x) \ge 0$, this random "Devil's staircase" (Mandelbrot 1983) is a non-decreasing function, as illustrated by the typical realization based on a log-normal cascade in Fig. A10a. The increments of (A40) are easily computed, using the definition of a coarse-grained measure in Eq. (14) in the main body:

$$f_\varepsilon(x+r)-f_\varepsilon(x) = \int_x^{x+r} \varepsilon(x')dx' = r\varepsilon(r;x),\ 0 \le x \le L-r,\ 0 \le r \le L. \tag{A41}$$

The qth order moment of this increment is directly related to that of $\varepsilon(r;x)$:

$$\langle |f_\varepsilon(x+r)-f_\varepsilon(x)|^q \rangle = r^q \langle \varepsilon(r;x)^q \rangle,\ q < q_D. \tag{A42}$$

Using the definitions in Eqs. (16) and (20) and taking logs, we find $\zeta(q) = q-K(q)$; equivalently,

$$H(q) = 1 - \frac{K(q)}{q},\ q < q_D. \tag{A43}$$

In particular, integrals of randomly placed δ-functions, with $K(q) = q-1$ ($q > 0$), yield randomly placed Heaviside steps which are multifractal functions with a broad range of Hölder exponents: $H(q) = 1/q$ ($q > 0$). The limit $H \to \infty$ in Eq. (A39) confirms this result, that Marshak *et al.* (1996) derive from first principles as well.

A.4.3 Fractionally Integrated Singular Cascades

To simulate the scale-invariant spatial properties observed in cloud fields, Schertzer and Lovejoy (1987) generalize the Devil's staircase concept. As a means of introducing the continuity that necessarily comes with nonstationarity, they simply use fractional integration instead of its standard counterpart. This yields what shall call a Fractionally Integrated Singular Cascade (FISC). We use here the same form of fractional integration as in §A.2.3, low-pass power-law filtering in Fourier space. Consider a 1D grid of size $M_n = 2^n$ and unitary outer scale,

$$x_m = j\ell_n,\ m = 0,\ldots,M_n-1\ (L = M_n\ell_n = 1), \tag{A44}$$

where we construct an n-step cascade process with branching ratio $\lambda = 2$. This field is generated in physical space and its Fourier representation is computed numerically:

$$\begin{cases} \varepsilon_n(x_j),\ j = 0,\ldots,2^n-1 \\ \tilde{\varepsilon}_n(k),\ k = 0,\pm1,\ldots,\pm2^{n-1}. \end{cases} \tag{A45}$$

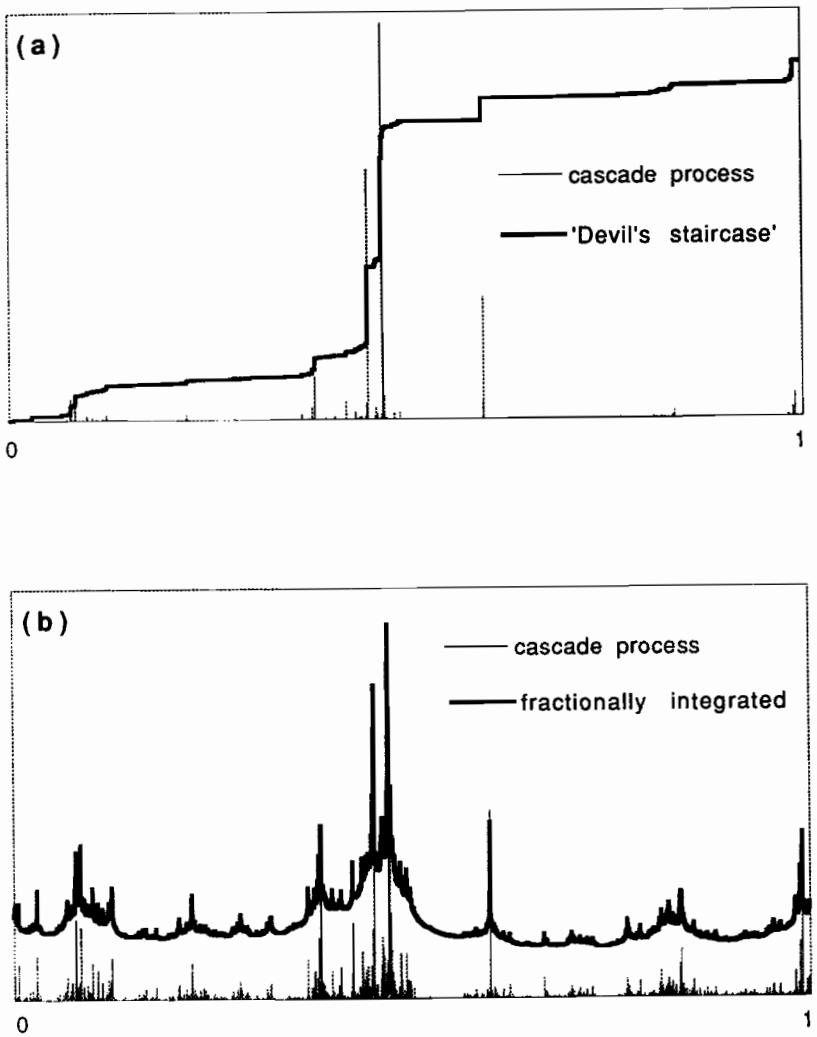


Figure A10: *Hybrid Multiplicative/Additive Models.* (*a*) Random Devil's Staircase: both the measure and its integral are illustrated. (*b*) Same as in panel (*a*) but for a *fractional* integration of order 1/3.

The measure $\varepsilon(x)$ is then "smoothed" into a function using

$$\begin{cases} f_n(x) = \varepsilon_n(x) * |x|^{-(1-H^*)} \\ \tilde{f}_n(k) = A(H^*)\, \tilde{\varepsilon}_n(k) \times |k|^{-H^*} \end{cases}, 0 < H^* < 1, \qquad \text{(A46}a\text{)}$$

where (e.g., Gradstein and Ryzhik 1980)

$$A(H^*) = \sqrt{\frac{2}{\pi}} \cos(\frac{\pi}{2} H^*) \Gamma(H^*). \qquad \text{(A46}b\text{)}$$

Figure A10*b* shows the outcome for a log-normal cascade similar to that in Fig. A10*a* using a fractional order of integration $H^* = 1/3$. The exponent H^* of the power-law filter is the nonstationarity parameter of the model, given directly by the spectral exponent we want for $f(x)$:

$$H^* = \frac{\beta_f - \beta_\varepsilon}{2}, \qquad \text{(A47)}$$

given that of $\varepsilon(x)$,

$$\beta_\varepsilon = 1 - K(2), \qquad \text{(A48)}$$

where $K(2)$ defines the scaling of $\langle \varepsilon(r;x)^2 \rangle \sim r^{-K(2)}$. As for fBm using the Fourier construction (§A.2.3), generalization to 2D is straightforward.

We need to estimate the scaling exponents for this model for $q \neq 2$ (at $q = 2$, the scaling is determined exactly by the construction algorithm). For this, we tentatively interpret Eq. (46*a*) as

$$|f(x+r) - f(x)| \sim \varepsilon(r;x)\, r^{H^*}; \qquad \text{(A49)}$$

taking qth powers, and averaging yields $\zeta(q) = qH^* - K(q)$, given the exponent definitions in main text's Eqs. (16) and (20). This $\zeta(q) \leftrightarrow K(q)$ connection is a special case of Eq. (26) with $a = 1/H^*$ and $b = 1$, leading to

$$H(q) = H^* - \frac{K(q)}{q}, q < q_D. \qquad \text{(A50)}$$

So this generalizes Eq. (A43) for Devil's staircases where $H^* = 1$.

However, the general (event-wise) applicability of (A49), hence (A50) for all values of q, to FISCs is questionable. However, Eq. (A50) is guaranteed by construction to work for $q = 0$ and $q = 2$, between these two values it provides at least a good approximation to numerically obtained $\zeta(q)$'s. As q increases far beyond 2, we are effectively emphasizing ever larger events in $\varepsilon(r;x)$ on the r.h.s. of (A49). It is easy to imagine situations where the increment $f(x+r) - f(x)$ on the l.h.s. of (A49) is small, although the underlying measure $\varepsilon(r;x)$ is large. For instance, take $x+r$ and x on either side of a strong spike in Fig. A10*b*. Consequently, the agreement between the statistics independently determined on either side of Eq. (A49) deteriorates.[1]

[1] Davis *et al.* (1993, 1994*b*) describe a more general approach for characterizing the correlations between some nonstationary process $f(x)$, artificial or real, and the associated $\varepsilon(x)$ field.

Table A1 (begin): *Scale-Invariant Processes.* 2nd-order statistics and associated stationarity and continuity properties.

	Model	Sect.	Figure	Domains? $x\in[0,L)^d$, w, f or ε		Parameters? primary	secondary	Stationarity? ... *per se*[a] ($\beta<1$) ... of incs[b] ($\beta<3$)		
	exponents:			d				β		
I.	**Discrete model for $w(x) = f'(x)$, without correlations:**									
a	Bernoulli w.n.	A.1.1		1	$w=\pm s$	$s>0$		0	√	√
b	Bm on a grid	A.1.1		1	$f/s\in\mathbb{Z}$	$s>0$		2		√
II.	**Gaussian model for $w(x) = f'(x)$, without correlations:**									
a	Gaussian w.n.	A.1.1	A1*a*	1,2	$w\in\Re$	$\sigma>0$	μ	0	√	√
b	off-grid Bm	A.1.1	A1*b*	1	$f\in\Re$	$\sigma>0$	f_0	2		√
III.	**Non-Gaussian model for $w(x) = f'(x)$, without correlations:**									
a	Lévy w.n.	A.1.2	A2*a*	1,2	$w\in\Re$	$0<\alpha<2$	*a*, *c*	0	√	√
b	Lévy-flight	A.1.2	A2*b*	1	$f\in\Re$	$0<\alpha<2$	*c*, f_0	2		√
IV.	**Gaussian models for $f'(x)$ and $f(x)$, with correlations:**									
a	SGSN+	A.2.4		1,2	$f'\in\Re$	$-1<H_2<0$	σ	$1-2\lvert H_2\rvert$	√	√
b	fBm	A.2	A3	1,2	$f\in\Re$	$0<H_2<1$	σ	$2H_2+1$		√
V.	**Cascade models $\varepsilon(x)$ and $f(x)$, with microcanonical conservation:**									
a	*p*-model	A.3.4	A6–7	1,	$\varepsilon>0$	$0<p<1/2$		$1-\log_2[1+(1-2p)^2]$	√	√
	3*p*-model	A.3.5		2		$0<p_i<1/2$ ($i=1,2,3$)		$1-\log_2\langle W^2\rangle$, with $q=2$ in (A33).		
b	Bounded Cascades	A.4.1	A6–8 A9	1, 2	$f_-\le f\le f_+$ ($0<f_-<f_+<\infty$)	$H>0$	p, p_i ($i=1,2,3$)	$\min\{2H,1\}+1$		√
VI.	**Cascade models for $\varepsilon(x)$ and a "hybrid" one for $f(x)$, with canonical conservation:**									
a	Beta-models	A.3.1	A4*a*, A5*a*	1,2	$\varepsilon\ge 0$	$0<D_f<d$		$1-(d-D_f)$	√	√
a'	lognormal casc.	A.3.3	A4*b*, A5	1,2	$\varepsilon>0$	$\sigma_{\ln W}>0$		$1-\sigma_{\ln W}^2/\ln 2$	√	√
b	FISC*	A.4.3	A10*b*	1,2	$f>\vec{0}$	$H^*>(1-\beta_\varepsilon)/2$ p, D_0, or $\sigma_{\ln W}$		$2H^*+\beta_\varepsilon$		√

[a]Models designated with an "*a*" are stationary in the "broad" sense, where the autocorrelation function depends only on the separation: $\langle w(x+r)w(x)\rangle \propto \delta(r)$ (I–III); $\langle f'(x+r)f'(x)\rangle \propto r^{2H_2}$ (IV, with $-1/2<H_2<0$); $\langle \varepsilon(x+r)\varepsilon(x)\rangle \propto r^{-K(2)}$ (V–VI, with $0<K(2)=1-\beta<1$). These models are also stochastically discontinuous since, e.g., $\langle[w(x+r)-w(x)]^2\rangle = 0$ for $r\ge 0$.

[b]The associated models designated with a "*b*" are (broad-sense) nonstationary but stochastically continuous in the sense that, as $r\to 0$, $\langle[f(x+r)-f(x)]^2\rangle \propto r^{\zeta(2)}\to 0$ with $0<\zeta(2)=\beta-1<2$. In categories I–III, one can go from model "*a*" to model "*b*" in $d=1$ simply by taking a running sum; they are "additive" models.

+Scaling Gaussian Stationary Noise.

*Fractionally Integrated Singular Cascade.

Table A1 (end): *Scale-Invariant Processes.* Multifractal statistics and associated ergodicity properties.

Multi-scaling? $\zeta(q)$ & ... $K(q)$		Inter-mittency? ($K(q)\not\equiv 0$) for w, f or ε & ... for $\lvert\nabla\rvert$'s		Bi-fractal properties? (position in $q=1$ plane) $H_1(w, f$ or $\varepsilon)$	$C_1(w, f$ or $\varepsilon)$	$C_1(\lvert\nabla\rvert$'s)	Ergodicity Properties? (remarks)
no	N.A.	N.A.	no	0	N.A. (negative values)	0	trivial ergodicity
no	N.A.	N.A.	no	1/2	N.A. (" ")	0	large fluctuations in local 1-pt. statistics and realization-to-realization differences.
no	N.A.	N.A.	no	0	N.A. (" ")	0	10^3 events are generally enough to sample up to $\approx 3\sigma$'s (cf. "3σ" rule).
no	N.A.	N.A.	no	1/2	N.A. (" ")	0	cf. I*b*
no	N.A.	N.A.	no	0	N.A. (" ")	0	moments of order $q\ge\alpha$ are divergent.
yes	N.A.	N.A.	no	$\min\{1/\alpha,1\}$	N.A. (" ")	0	same as I*b* and II*b*, but worse due to divergences.
no	N.A.	N.A.	no	0	N.A. (" ")	0	cf. II*a*
no	N.A.	N.A.	no	H_2	N.A. (" ")	0	cf. II*b*
no	yes	yes	yes	0	$p\log_2 p+(1-p)\log_2(1-p)+1$ $(d/dq)\log_2\langle W^q\rangle\lvert_{q=1} = \langle W\log_2 W\rangle$, from (A33).	$C_1(\varepsilon)$#	$\langle\varepsilon(r;x)\rangle$ dsitribution is independent of the realization. $\langle\varepsilon(L;0)\rangle$ is independent of realization.
yes	no	no	N.A.†	$\min\{H,1\}$	0	N.A.†	although on a bounded domain, between $f_\pm = \prod_{i=0}^{\infty}[1\pm(1-2p)/2^{iH}]$, the pdf has lognormal-like skewness and the associated sampling problems; increments are also broadly distributed.
no	no	yes	yes	0	$d-D_f$	$C_1(\varepsilon)$#	no divergence in $\langle\varepsilon(r;x)^q\rangle$.
no	yes	yes	yes	0	$\sigma_{\ln W}^2/2\ln 2$	$C_1(\varepsilon)$#	$\langle\varepsilon(r;x)^q\rangle$ divergenes for $q\ge d/C_1$.
yes	no	no	yes	H^*	0	$C_{1\varepsilon}$§	$K(q)$ for $\lvert\nabla f\rvert$'s: cf. corresponding ε-model.

†In standard singularity analysis using next-neighbor absolute differences, spurious scale-breaks occur due to the binary-tree construction of this model.

#Taking absolute gradients of a singular cascade model only emphasizes the spikes, and the same singularity spectrum is found (Lavallée *et al.*, 1993).

§Absolute gradients of a FISC have the same singularity properties as the original cascade (Lavallée *et al.*, *ibid.*).

A.5 Summary

We have described specific algorithms for constructing mono- and multiscaling stationary measures amenable to singularity analysis (Section 4.4 in main text), as well as mono- and multiscaling nonstationary functions with stationary increments amenable to structure function analysis (Section 4.2 in main body). In each case, we have specified the dependency of the appropriate scaling exponents on the parameters of the model. One example in two spatial dimensions is described explicitly and, in all others, generalization from 1D to 2D is straightforward. This collection of models is comprehensive enough to calibrate and study the sensitivity of any standard multifractal data analysis procedure using arbitrary amounts of synthetic data with controllable statistical properties. In particular, convergence rates of spatial averages to their ensemble counterparts can be investigated since the models have variable degrees of ergodicity. Table A1 organizes the models by category, lists their parameters and summarizes their properties, while Fig. A11 displays their inter-relations graphically.

Validation of analysis procedures is only one application for stochastic modeling. In our specific area of research, the effects of internal cloud structure on radiative properties, we have used multifractal models as artificial clouds-in-a-computer with controllable properties. Extensive numerical experimentation, using Monte Carlo and other radiative transfer techniques, has lead us to new insight into the ways clouds affect the Earth's radiative budget (Cahalan 1994) and ways of retrieving cloud properties from satellite (Marshak *et al.* 1995*a,b*) and lidar (Davis *et al.* 1996*b*) data.

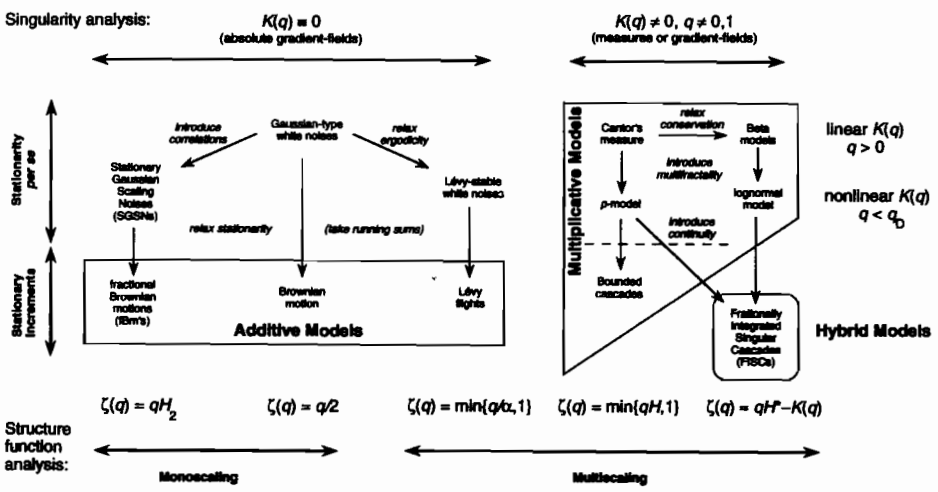


Figure A11: *Classification of Scale-Invariant Models.* The formulas for the qth-order structure function exponents refer to the models in the last row that are nonstationary but with stationary increments.

Acknowledgments

This work was supported by the Environmental Sciences Division of U.S. Department of Energy (under grant DE-A105-90ER61069 to NASA's Goddard Space Flight Center) as part of the Atmospheric Radiation Measurement (ARM) program. We thank A. Arnéodo, T. Bell, S. Lovejoy, Ch. Meneveau, D. Schertzer, and T. Warn for helpful discussions.

References

Arnéodo, A., Grasseau, G., and Holschneider, M., "Wavelet transform of multifractals," *Phys. Rev. Lett.* **61** (1988) 2281–2284.

Arnéodo, A., J. F. Muzy, and E. Bacry, "Wavelet analysis of fractal signals. Applications to fully developed turbulence data," in *Eddy Structure Identification in Free Turbulent Shear Flows*, eds. J.P. Bonnet and M.N. Glauser (1992), p. 153.

Aurell, E., U. Frisch, J. Lutsko, and M. Vergassola, "On the multifractal properties of the energy dissipation derived from turbulence data," *J. Fluid Mech.* **238** (1992) 467–486.

Benzi, R., L. Biferale, A. Crisanti, G. Paladin, M. Vergassola, and A. Vulpani, "A random process for the construction of multi-affine fields," *Physica D* **65** (1993) 352–358.

Bertozzi, A. L, and A. B. Chhabra, "Cancellation exponents and fractal scaling," *Phys. Rev. E* **49** (1994) 4716–4719.

Cahalan, R. F., M. Nestler, W. Ridgway, W. J. Wiscombe, and T. L. Bell, "Marine stratocumulus spatial structure," in *Proceedings of the 4th International Meeting on Statistical Climatology*, ed. J. Sansom (New Zealand Meteorological Service, Wellington, 1990), pp. 28-32.

Cahalan, R. F., W. Ridgway, W. J. Wiscombe, T. L. Bell, and J. B. Snider, "The albedo of fractal stratocumulus clouds," *J. Atmos. Sci.* **51** (1994) 2434–2455.

Cahalan, R. F., "Bounded cascade clouds: albedo and effective thickness," *Nonlin. Proc. in Geophysics* **1** (1994) 156–167.

Christakos, G., *Random Fields in Earth Sciences* (Academic Press, San Diego, 1992).

Coakley, J. A. (Jr.), and F. R. Bretherton, "Cloud cover from high-resolution scanner data: Detecting and allowing for partially filled fields of view," *J. Geophys. Res.* **87** (1982) 4917–4932.

Corrsin, S., "On the spectrum of isotropic temperature fluctuations in isotropic turbulence," *J. Appl. Phys.* **22** (1951) 469–473.

Davis, A., A. Marshak, and W. Wiscombe, "Bi-multifractal analysis and multi-affine modeling of nonstationary geophysical processes – application to turbulence and clouds," *Fractals* **1** (1993) 560–567.

Davis, A., A. Marshak, W. Wiscombe, and R. Cahalan, "Multifractal characterizations of nonstationarity and intermittency in geophysical fields, observed, retrieved or simulated," *J. Geophys. Res.* **99** (1994*a*) 8055–8072.

Davis, A., A. Marshak, and W. Wiscombe, "Wavelet-based multifractal analysis of nonstationary and/or intermittent geophysical signals," in *Wavelets in Geophysics*, eds. E. Foufoula-Georgiou and P. Kumar (Academic Press, San Diego, 1994*b*), pp. 249–298.

Davis, A., A. Marshak, W. Wiscombe, and R. Cahalan, "Scale-invariance in liquid water distributions in marine stratocumulus – Part I: Spectral properties and stationarity issues," *J. Atmos. Sci.* (1996*a*, in press).

Davis, A., A. Marshak, R. Cahalan, and W. Wiscombe, "Horizontal radiative fluxes in stratocumulus and the landsat scale-break," *J. Atmos. Sci.* (1996*b*, accepted).

Davis, A., and A. Marshak, "Bi-Fractal Analysis of Atmospheric SignalsIllustrated with Cloud Data," in *Proceedings of the 1995 Battlefield Atmospheric Conference* (1996, in press).

Eneva, M., "Fractal description of microseismic activity," *Nonlin. Proc. in Geophysics* **1** (1994) 65–76.

Evertsz, C. J. G., and B. B. Mandelbrot, "Multifractal measures," in *Chaos and Fractals, New Frontiers of Science*, eds. H.-O. Peitgen, H. Jürgens and D. Saupe (Springer-Verlag, New York, 1992), pp. 921–953.

Falconer, K. J., *Fractal Geometry - Mathematical Foundations and Applications*, pp.xxii+288, J. Wiley, New York, 1990.

Feigenbaum, M. J., "Some characterizations of strange sets," *J. Stat. Phys.* **46** (1987) 919–924.

Feller, W., *An Introduction to Probability Theory and its Applications*, vol. 2, (Wiley, New York, 1971).

Frisch, U., P. L. Sulem, and M. Nelkin, "A simple dynamical model of intermittent fully developed turbulence," *J. Fluid Mech.* **87** (1978) 719–724.

Frisch, U., "From global scaling, *à la* Kolmogorov, to local multifractality in fully developed turbulence," *Proc. R. Soc. Lond. A* **434** (1991) 89–99.

Gerber, H., B. G. Arends, and A. S. Ackerman, "New microphysics sensor for aircraft use," *Atmos. Res.* **31** (1994) 235-252.

Gradshteyn, I. S., and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic Press, San Diego, 1980).

Grassberger, P., "Generalized dimensions of strange attractors," *Phys Rev. Lett.* **97** (1983) 227–.

Grivet-Talocia, S., "On the scale dependence of fractal dimension for band-limited $1/f^{\alpha}$ noise," *Physics Lett. A* **200** (1995) 264–276,

Gupta, V. K., and E. C. Waymire, "A statistical analysis of mesoscale rainfall as a random cascade," *J. Appl. Meteor.* **32** (1993) 251–267.

Halsey, T. C., M. H. Jensen, L. P. Kadanoff, I. Procaccia, and B. Shraiman, "Fractal measures and their singularities – The characterization of strange sets," *Phys. Rev. A* **33** (1986) 1141–1151.

Hentschel, H. G. E., and I. Procaccia, "The infinite number of generalized dimensions of fractals and strange attractors," *Physica D* **8** (1983) 435–444.

Jensen, J. L. W. V., "Sur les fonctions convexes et les inégalités entre les valeurs moyennes," *Acta Math.* **30** (1906) 789–806.

Kahane, J. P., and J. Perièyre, "Sur certaines martingales de Benoit Mandelbrot," *Adv. in Math.* **22** (1976) 131-145.

Kolmogorov, A. N., "Local structure of turbulence in an incompressible liquid for very large reynolds numbers," *Dokl. Akad. Nauk SSSR* **30** (1941) 299–303.

Kolmogorov, A. N., "A Refinement of Previous Hypothesis Concerning the Local Structure of Turbulence in Viscous Incompressible Fluid at High Reynolds Number," *J. Fluid Mech.* **13** (1962) 82–85.

Lavallée, D., S. Lovejoy, D. Schertzer, and P. Ladoy, "Nonlinear variability and simulation of landscape topography,"in *Fractals in Geography*, eds. L. De Cola and M. Lam (Kluwer, Boston, 1993) pp. 158–192.

Mallat, S. G., "A theory for multiresolution signal decomposition: the wavelet transformation," *IEEE Trans. Pattern Anal. Mach. Intel.* **11** (1989) 674–693.

Marshak A., A. Davis, R. Cahalan, and W. Wiscombe, "Bounded cascade models as nonstationary multifractals," *Phys. Rev. E* **49** (1994) 55-69.

Marshak A., A. Davis, W. Wiscombe, and G. Titov, "The verisimilitude of the independent pixel approximation used in cloud remote sensing," *Remote Sens. Environ.* **52** (1995*a*) 71–78.

Marshak A., A. Davis, and W. Wiscombe, "Radiation smoothing in fractal clouds." *J. Geophys. Res.* **100** (1995*b*) 26247–26261.

Marshak, A., A. Davis, W. J. Wiscombe, and R. F. Cahalan, "Scale-invariance of liquid water distributions in marine stratocumulus – Part 2: Multifractal properties and intermittency issues," *J. Atmos. Sci.* (1996, submitted).

Mandelbrot, B. B., "Intermittant turbulence and self-similar cascades: Divergence of high moments and dimension of the carrier," *J. Fluid Mech.* **62** (1974) 331–350.

Mandelbrot, B. B., *The Fractal Geometry of Nature* (Freemann, New York, 1983).

Mandelbrot, B. B., and J. W. van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Review* **10** (1968) 422-437.

Meneveau, C., and K. R. Sreenivasan, "The multifractal spectrum of the dissipation field in turbulent flows," *Nuclear Phys. B* **2** (1987*a*) 49-76.

Meneveau, C., and K. R. Sreenivasan, "Simple multifractal cascade model for fully developed turbulence," *Phys. Review Lett.* **59** (1987*b*) 1424–1427.

Meneveau, C., and K. R. Sreenivasan, "Multifractal nature of turbulent energy dissipation," *J. Fluid Mech.* **224** (1991) 429–.

Monin, A. S., and A. M. Yaglom, *Statistical Fluid Mechanics*, vol. 2 (MIT Press, Boston, 1975).

Muzy, J.-F., E. Bacry, and A. Arnéodo, "Multifractal formalism for fractal signals: The structure-function approach versus the wavelet-transform modulus-maxima method," *Phys. Rev. E* **47** (1993) 875–884.

Muzy, J.-F., E. Bacry, and A. Arnéodo, "The multifractal formalism revisited with wavelets." *Int. J. of Bifurcation and Chaos*, **4** (1994), 245–302.

Novikov, E. A., and R. Stewart, "Intermittancy of turbulence and spectrum of fluctuations in energy dissipation," *Izv. Akad. Naut. SSSR, Ser. Geofiz.* **3** (1964) 408–412.

Obukhov, A., "Structure of the temperature field in a turbulent flow," *Izv. Akad. Naut. SSSR, Ser. Geogr. I Geofiz.* **13** (1949) 55–69.

Obukhov, A., "Some specific features of atmospheric turbulence," *J. Geophys. Res.* **67** (1962) 3011–3014.

Ott, E., Y. Du, K.R. Sreenivasan, A. Juneja, and A.K. Suri, "Sign-singular measures: Fast magnetic dynamos, and high Reynolds-number fluid turbulence. *Phys. Rev. Lett.* **69** (1992) 2654–2657.

Parisi, G., and U. Frisch, "A multifractal model of intermittency," in *Turbulence and Predicability in Geophysical Fluid Dynamics and Climate Dynamics*, eds. M. Ghil, R. Benzi, and G. Parisi (North-Holland, 1985), pp. 84–88.

Pearson, C. E. (ed.), *Handbook of Applied Mathematics – Selected Results and Methods*, 2nd ed. (Van Nostrand Reinhold, New York, 1990).

Peitgen, H.-O., and D. Saupe (eds.), *The Science of Fractal Images* (Springer-Verlag, 1988).

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN*, 2nd ed. (Cambridge University Press, New York, 1993).

Schertzer, D., and S. Lovejoy, "Physical modeling and analysis of rain clouds by anisotropic scaling multiplicative processes. *J. Geophys. Res.* **92** (1987) 9693–9714.

Schertzer D., and S. Lovejoy, "Hard and soft multifractal processes," *Physica A* **185** (1992) 187–194.

Sreenivasan, K. R., "Fractals and multifractals in fluid turbulence," *Ann. Rev. Fluid. Mech.* **23** (1991) 539–600.

Sykes, R. I., R. S. Gabruk, and D. S. Henn, "A fractal/multifractal representation of the small-scale structure in a turbulent plume," *J. Appl. Meteor.* (1996, in press).

Vainshtein, S. I., K. R. Sreenivasan, R. T. Pierrehumbert, V. Kashyap, and A. Juneja, "Scaling exponents for turbulence and other random processes and their relationships with multifractal structure, *Phys. Rev. E* **50** (1994) 1823-1835.

Viscek, T., and A.-L. Barabási, "Multi-affine model for the velocity distribution in fully turbulent flows," *J. Phys. A: Math. Gen.*, **24** (1991) L845–L851.

Voss, R. F., "Fourier synthesis of Gausian fractals: 1/f noises, landscapes, and flakes," in *Proc. Siggraph Conf.* (Detroit, 1983), pp. 1–21.

Waymire, E., and V. J. Gupta, "The mathematical structure of rainfall representations, Parts 1–3," *Water Resour. Res.* **17** (1981) 1261–1294.

Wiscombe, W., A. Davis, and A. Marshak, "Scale-invariance, nonstationarity and intermittency in the structure of cloud cover," in *Proceedings of the 4th Atmospheric Radiation Measurement (ARM) Science Team Meeting* (U.S. Dept. of. Energy, Washington DC, 1995), pp. 11–14.

Current Topics in Nonstationary Analysis

Proceedings of the Second Workshop on Nonstationary Random Processes and their Applications

San Diego, CA June 11 – 12, 1995

edited by

George Treviño
CHIRES Associates, Inc.

Jay Hardin
Nasa-Langley Research Center

Bruce Douglas
Naval Surface Warfare Center-Carderock

Edgar Andreas
US Army Cold Regions Research and Engineering Laboratory

World Scientific
Singapore • New Jersey • London • Hong Kong

Published by

World Scientific Publishing Co. Pte. Ltd.

P O Box 128, Farrer Road, Singapore 912805

USA office: Suite 1B, 1060 Main Street, River Edge, NJ 07661

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

CURRENT TOPICS IN NONSTATIONARY ANALYSIS

Proceedings of the Second Workshop on Nonstationary Random Processes and their Applications

ISBN 981-02-2703-5

Printed in Singapore.

PREFACE

Nonstationarity is another name for intermittency, a phenomenon which affects many physical processes in the atmospheric boundary layer. Among these are the transfers of heat, momentum, and moisture, and thus the propagation of electromagnetic waves. Nonstationarity also appears in the nonlinear propagation of acoustic waves and in noise radiation by supersonic jets and helicopters. Such signals are commonplace in military communication systems, ship and submarine stealth, and maneuvering and control problems. Also, data collected in many diverse R&D programs frequently exhibit nonstationary features. Of these a common one is turbulence data, which we know from observations to be a complicated nonlinear dynamical phenomenon of limited predictability.

Several advances in the theory of nonstationary random processes have been made since the first workshop convened in 1991. And although much research remains yet to be done, it seems opportune to convert some of the existing work into a more permanent form. It is debatable what impact the new knowledge has had on the development of practical models of this important non-equilibrium phenomenon. The chief aim of the workshop was therefore to provide a forum at which the recent important contributions could be reported and discussed among researchers from government, academia, and industry.

The theme of the workshop was all aspects of nonstationary analysis, with the appeal for participation being made to engineers, scientists, and mathematicians alike. This appeal is consistent with that of the first workshop. The intent was to create a diverse environment for researchers working in this genuinely multidisciplinary field to mutually share their ideas. The premier objective of the workshop was to consolidate recent developments in nonstationary analysis and present the material in a "tutorial mode." A second objective was to delineate open problems.

We would like to express our gratitude to those who assisted us in convening the workshop. Sincere thanks go to Jack Preisser (NASA Langley), Frank Halsall (NSWC-Carderock), and Walter Bach (Army Research Office), who collectively provided most of the financial support necessary to bring the workshop to fruition. We gratefully acknowledge the administrative support provided by Norma Treviño. To the anonymous referees who reviewed the submitted manuscripts, and to those who served as session chairmen (Frank Halsall, Alan Piersol and Ken Bolland), we are also grateful. Most of all, though, we express our deep appreciation to the authors, whose hard work and dedication made the workshop a success.

George Treviño
Jay Hardin
Bruce Douglas
Edgar Andreas